

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:30:38 ; Search time 51.57 seconds
(without alignments)
5.314 Million cell updates/sec

Title: PCT-US01-05825A-30
Perfect score: 43
Sequence: 1 VDEGRGRC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 segs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_39:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	334	1	G3P1_HUMAN
2	38	88.4	25	1	P800354 homo sapien
3	38	88.4	320	1	P80447 jaculus orl
4	38	88.4	332	1	G3P2_BOVIN
5	38	88.4	332	1	G3P_CHICK
6	38	88.4	332	1	G3P_COILI
7	38	88.4	332	1	G3P_COTJA
8	38	88.4	332	1	G3P_CRIGR
9	38	88.4	332	1	G3P_MOUSE
10	38	88.4	332	1	G3P_PIG
11	38	88.4	332	1	G3P_RABIT
12	38	88.4	333	1	G3P_RAT
13	38	88.4	334	1	G3P2_ZYGRO
14	38	88.4	334	1	G3P1_HUMAN
15	37	86.0	35	1	G3P1_JACOR
16	37	86.0	303	1	G3P1_ANASP
17	37	86.0	324	1	G3P1_KLEPN
18	37	86.0	329	1	G3P1_GLORO
19	37	86.0	330	1	G3P1_KLULA
20	37	86.0	330	1	G3P1_ECOLI
21	37	86.0	330	1	G3P3_LEIME
22	37	86.0	331	1	G3P3_TRYBB
23	37	86.0	331	1	G3P1_YEAST
24	37	86.0	331	1	G3P2_KLUMA
25	37	86.0	331	1	G3P2_YEAST
26	37	86.0	331	1	G3P3_YEAST
27	37	86.0	331	1	G3P3_CANAL
28	37	86.0	331	1	G3P_MONAN
29	37	86.0	332	1	G3P1_THEAO
30	37	86.0	332	1	G3P1_DROME
31	37	86.0	332	1	G3P2_DROME
32	37	86.0	332	1	G3P3_ECOLI
33	37	86.0	332	1	G3P_BUCAI
34	37	86.0	332	1	G3P_BUCAP

34	37	86.0	332	1	G3P_DROME	001597 drosophila
35	37	86.0	332	1	G3P_HELPJ	092410 helicobacte
36	37	86.0	332	1	G3P_HELPJ	P55571 helicobacte
37	37	86.0	332	1	G3P_PHYIN	P26988 phytophthor
38	37	86.0	332	1	G3P_STRAU	059800 streptomyc
39	37	86.0	332	1	G3P_THEMA	P17721 thermotoga
40	37	86.0	333	1	G3P2_RHOSH	P29272 rhodobacter
41	37	86.0	333	1	G3P_HOMAM	P00357 homarus ame
42	37	86.0	333	1	G3P_PALVE	P56649 palinurus v
43	37	86.0	333	1	G3P_PICPA	092263 pichia past
44	37	86.0	333	1	G3P_PICPA	P54226 streptomyce
45	37	86.0	334	1	G3P1_ANAVA	P34916 anabaena va

ALIGNMENTS

RESULT	1	STANDARD	PRT	334 AA.
ID	G3P1_HUMAN			
AC	P00354;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, MUSCLE (EC 1.2.1.12).			
GN	GAPD.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=82073291; PubMed=7030790;			
RA	Nowak K., Wolny M., Banas T.;			
RT	"The complete amino acid sequence of human muscle glyceraldehyde 3-			
RT	phosphate dehydrogenase.";			
RL	FEBS Lett. 134:143-146(1981).			
RN	[2]			
RP	PARTIAL SEQUENCE.			
RX	MEDLINE=76067491; PubMed=1193541;			
RA	Nowak K., Kuczek M., Ostropolska L., Malarska A., Wolny M.,			
RA	Branowski T.;			
RT	"The covalent structure of glyceraldehyde-phosphate dehydrogenase			
RT	from human muscles. Isolation and amino acid sequences of peptides			
RT	from tryptic digest.";			
RL	Hoppe-Seyler's Z. Physiol. Chem. 356:1181-1183(1975).			
RN	[3]			
RP	X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).			
RX	MEDLINE=76265083; PubMed=957435;			
RA	Merger W.D., Wilm S.I., Watson H.C.;			
RT	"Twining in crystals of human skeletal muscle D-glyceraldehyde-3-			
RT	phosphate dehydrogenase.";			
RL	J. Mol. Biol. 104:277-283(1976).			
CC	-1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE			
CC	+ NAD(+) -> 1,3-DIPHOSPHATEGLYCERATE + NADH.			
CC	-1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.			
CC	-1- SUBUNIT: HOMOTETRAMER.			
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC.			
CC	-1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE			
CC	DEHYDROGENASE FAMILY.			
DR	PIR; A00366; DEHUG3.			
DR	PDB; 3GPD; 16-JUL-88.			
DR	MIM; 138400; -.			
DR	InterPro; IPR000173; -.			
DR	Pfam; PF00044; gpdh; 1.			
DR	PRINTS; PR00078; G3PDHGNASE.			
DR	PROSITE; PS00071; GAPDH; 1.			
KW	GLYCOLYSIS; Oxidoreductase; NAD, Multigene family; 3D-structure.			
FT	BINDING	151	151	
FT	ACT. SITE	178	178	
FT	STRAND	5	8	
FT	TURN	12	12	
FT	HELIX	13	23	

```

FT TURN 24 25
FT STRAND 32 33
FT HELIX 39 47
FT STRAND 48 49
FT TURN 50 52
FT STRAND 53 54
FT STRAND 59 60
FT STRAND 65 68
FT TURN 69 70
FT STRAND 71 76
FT HELIX 81 83
FT TURN 87 90
FT STRAND 93 96
FT HELIX 104 113
FT STRAND 117 121
FT TURN 130 130
FT TURN 132 134
FT HELIX 136 138
FT TURN 141 142
FT STRAND 145 148
FT TURN 151 152
FT HELIX 153 167
FT STRAND 170 178
FT STRAND 188 188
FT TURN 193 194
FT TURN 197 198
FT STRAND 199 199
FT STRAND 206 209
FT HELIX 212 219
FT STRAND 227 233
FT STRAND 241 247
FT HELIX 254 266
FT TURN 267 269
FT TURN 282 283
FT TURN 298 299
FT STRAND 301 303
FT TURN 304 305
FT STRAND 306 310
FT HELIX 315 333
SQ SEQUENCE 334 AA; 35876 MM; 0C67E5CD6CDC292D CRC64;

```

```

Query Match      100.0%; Score 43; DB 1; Length 334;
Best Local Similarity 100.0%; Pred. NO. 0.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VDGFRIG 8
DB 7 VDGFRIG 14

```

```

RESULT 2
G3P2_JACOR STANDARD; PRT; 25 AA.
AC P80447;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, LIVER (EC 1.2.1.12) (GAPDH)
DE (FRAGMENT).
OS GAPD OR GAPC.
GN Jaculus orientalis.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Dipodidae; Dipodinae;
OC Jaculus.
OX NCBI_TaxID=48868;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=96139342; PubMed=8547342;
RA Soukri A., Hafid N., Valverde F., Elkeba M.S., Serrano A.;
RT "Evidence for a posttranslational covalent modification of liver
glyceraldehyde-3-phosphate dehydrogenase in hibernating jerboa

```

```

RT (Jaculus orientalis).";
RL Biochim. Biophys. Acta 1292:177-187(1996).
CC -I- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC + NAD(+) = 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC -I- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -I- SUBUNIT: HOMOTETRAMER.
CC -I- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -I- PTM: THE HIBERNATING ADULT ISOFORM IS ADP-RIBOSYLATED. THIS
CC ACCOUNTS FOR 2-3 FOLD LOWER SPECIFIC ACTIVITY IN THE HIBERNATING
CC ISOFORM.
CC -I- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
CC HSSP; P00354; 3GPD.
DR InterPro; IPR000173; -.
DR PROSITE; PS00071; GAPDH: PARTIAL.
KW Glycolysis; Oxidoreductase; NAD, ADP-ribosylation.
FT VARIANT 6 6 N->D (IN HIBERNATING ADULT LIVER
FT ISOFORM).
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 2603 MM; 4FC96C356CE79F2D CRC64;

```

```

Query Match      88.4%; Score 38; DB 1; Length 25;
Best Local Similarity 87.5%; Pred. NO. 0.62;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 VDGFRIG 8
DB 5 VDGFRIG 12

```

```

RESULT 3
G3P_BOVIN STANDARD; PRT; 320 AA.
ID G3P_BOVIN
AC P10096; P79130;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12) (GAPDH)
DE (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE OF 1-30.
RC TISSUE=Liver;
RX MEDLINE=76087882; PubMed=1201027;
RA Kuibe K., Jackson K.W., Tang J.;
RT "Structural evidence for a liver-specific glyceraldehyde-3-phosphate
dehydrogenase."
RL Biochem. Biophys. Res. Commun. 67:35-42(1975).
RN [2]
RP SEQUENCE OF 10-320 FROM N.A.
RC TISSUE=Lymphocytes;
RA Mertens B., Murtuki C.;
RT Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC + NAD(+) = 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC -I- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -I- SUBUNIT: HOMOTETRAMER.
CC -I- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -I- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.

```

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)


```

CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL; U85042; AAB47507.1; -.
CC PIR; A12055; A12055.
CC HSSP; P00354; 3GPD.
CC InterPro; IPR000173; -.
CC Pfam; PF00044; gpdb; 1.
CC PROSITE; PS00071; GAPDH; 1.
CC GlycoLysis; Oxidoreductase; NAD.
CC BINDING 149 149 GLYCERALDEHYDE 3-PHOSPHATE.
CC ACT_SITE 176 176 ACTIVATES THIOL GROUP DURING CATALYSIS.
CC CONFLICT 29 29 A -> F (IN REF. 1).
CC NON_TER 320 320
CC SEQUENCE 320 AA; 34382 MW; 92F0AF3C6C504F5B CRC64;

Query Match 88.4%; Score 38; DB 1; Length 320;
Best Local Similarity 87.5%; Pred. No. 6.1;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDFGCRIG 8
Db 5 VNGFCRIG 12

RESULT 4
G3P_CHICK STANDARD; PRT; 332 AA.
ID G3P_CHICK
AC P00356; Q90848; Q90826; Q90849;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12) (GAPDH).
GN GAPD.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI-TaxID=9031;
CX [1]
RN SEQUENCE FROM N.A.
RA MEDLINE=84153854; PubMed=6322764;
RA Parabierles F., Plechaczkyk M., Rainer B., Dani C., Fort P., Rlaad S.,
RA Maty L., Imbach J.L., Jeanneur P., Blanchard J.M.;
RT "Complete nucleotide sequence of the messenger RNA coding for chicken
RL muscle glyceraldehyde-3-phosphate dehydrogenase."
RN [2]
RN Biochem. Biophys. Res. Commun. 118:767-773(1984).
RN SEQUENCE FROM N.A.
RA MEDLINE=83204759; PubMed=6303388;
RA Dugalczyk A., Hazon J.A., Stone E.M., Dennison O.E., Rothblum K.N.,
RA Schwartz R.J.;
RT "Cloning and sequencing of a deoxyribonucleic acid copy of
RT glyceraldehyde-3-phosphate dehydrogenase messenger ribonucleic acid
RT isolated from chicken muscle."
RN [3]
RN Biochemistry 22:1605-1613(1983).
RN SEQUENCE FROM N.A.
RA MEDLINE=8516184; PubMed=3856841;
RA Stone E.M., Rothblum K.N., Alevy M.C., Kuo T.M., Schwartz R.J.;
RT "Complete sequence of the chicken glyceraldehyde-3-phosphate
RT dehydrogenase gene."
RN [4]
RN Proc. Natl. Acad. Sci. U.S.A. 82:1628-1632(1985).
RN SEQUENCE OF 196-332 FROM N.A.
RC TISSUE-Heart;
RA MEDLINE=82265644; PubMed=6179937;
RA Arnold H.H., Domdey H., Wiebauer K., Datta K., Siddiqui M.A.Q.;
RT "Cloning, partial sequencing, and expression of glyceraldehyde-3-
RT phosphate dehydrogenase gene in chick embryonic heart muscle cells."
RN [5]
RN J. Biol. Chem. 257:9872-9877(1982).
RN REVISION TO 328.

```

```

RA Arnold H.H., Domdey H., Wiebauer K., Datta K., Siddiqui M.A.Q.;
RL J. Biol. Chem. 258:2063-2063(1983).
RN [6]
RN SEQUENCE OF 1-224 FROM N.A.
RP TISSUE-Brain;
RC MEDLINE=83220743; PubMed=6687938;
RA Milner R.J., Brow M.A.D., Cleveland D.W., Shinnick T.M.,
RA Sutcliffe J.G.;
RT "Glyceraldehyde 3-phosphate dehydrogenase protein and mRNA are both
RT differentially expressed in adult chickens but not chick embryos."
RL Nucleic Acids Res. 11:3301-3315(1983).
CC -1- GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC + NAD(+) - 1,3-DIPHOSPHATEGYCERATE + NADH.
CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL; V00407; CAA23698.1; -.
CC EMBL; K01458; AAA48778.1; -.
CC EMBL; M1213; AAA48774.1; -.
CC EMBL; V00406; CAA23697.1; -.
CC EMBL; J00848; AAA48777.1; -.
CC EMBL; X01578; CAA25733.1; -.
CC PIR; A00368; DECHG3.
CC PIR; A22035; A22035.
CC PIR; A32737; A32737.
CC HSSP; P00354; 3GPD.
CC InterPro; IPR000173; -.
CC Pfam; PF00044; gpdb; 1.
CC PRINTS; PR00078; G3PDHDEGNASE.
CC PROSITE; PS00071; GAPDH; 1.
CC GlycoLysis; Oxidoreductase; NAD.
CC INIT MET 0 0
CC ACT_SITE 149 149 GLYCERALDEHYDE 3-PHOSPHATE.
CC CONFLICT 176 176 ACTIVATES THIOL GROUP DURING CATALYSIS.
CC CONFLICT 1 1 V -> RSE (IN REF. 6).
CC CONFLICT 144 144 V -> L (IN REF. 6).
CC CONFLICT 196 196 G -> D (IN REF. 1, 2 AND 3).
CC CONFLICT 276 276 D -> E (IN REF. 4).
CC CONFLICT 293 293 D -> H (IN REF. 3).
CC CONFLICT 328 328 M -> T (IN CAA23687).
CC SEQUENCE 332 AA; 35573 MW; 3CF1501058A9E9A0 CRC64;

Query Match 88.4%; Score 38; DB 1; Length 332;
Best Local Similarity 87.5%; Pred. No. 6.3;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDFGCRIG 8
Db 5 VNGFCRIG 12

RESULT 5
G3P_COLLIT STANDARD; PRT; 332 AA.
ID G3P_COLLIT
AC O57479;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12) (GAPDH).
GN GAPD OR GAPDH.
OS Columba livia (Domestic pigeon).

```

```

CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Archosauria; Aves; Neognathae; Columbiformes; Columba;
CC RN NCBI_TaxID=8932;
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Liver;
CC RA Mews J.G., Lambeth D.O.;
CC RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC CC + NAD(+) -> 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC CC DEHYDROGENASE FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL: AF036934; AAB8869.1; -
CC DR InterPro: IPR000173; -
CC DR Pfam: PF00044; gpdh; 1.
CC DR PRINTS: PR00078; G3PDHRCNASE.
CC DR PROSITE: PS00071; GAPDH: 1.
CC KW Glycolysis; Oxidoreductase; NAD.
CC FT INIT_MET 0 BY SIMILARITY.
CC FT BINDING 149 149 GLYCERALDEHYDE 3-PHOSPHATE.
CC FT ACT_SITE 176 176 ACTIVATES THIOL GROUP DURING CATALYSIS.
CC SQ SEQUENCE 332 AA; 35636 MW; 2DC7ACB75C2DEB24 CRC64;

Query Match 88.4%; Score 38; DB 1; Length 332;
Best Local Similarity 87.5%; Pred. No. 6.3;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDFGFRIG 8
Db 5 VNGFGRIG 12

RESULT 6
G3P_COTVA STANDARD; PRT; 332 AA.
AC 005025;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12) (GAPDH).
GN GAPD.
OS Coturnix coturnix japonica (Japanese quail).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC OC Coturnix.
CC RN NCBI_TaxID=93934;
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Fibroblast;
CC RA MEDLINE-93329297; PubMed-8514192;
CC RT Weisskirchen R., Stenmister G., Hartl M., Bister K.;
CC RL "Sequence and expression of a glyceraldehyde-3-phosphate
CC RL dehydrogenase-encoding gene from quail embryo fibroblasts.";
CC RL Gene 128:269-272(1993).
CC CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC CC + NAD(+) -> 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC CC -1- SUBUNIT: HOMOTETRAMER.
CC CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE

```

```

CC CC DEHYDROGENASE FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL: Z19086; CAA79512.1; -
CC DR PIR: S35726; S35726.
CC DR PIR: JN0678; JN0678.
CC DR HSP: P00354; 3GPD.
CC DR InterPro: IPR000173; -
CC DR Pfam: PF00044; gpdh; 1.
CC DR PRINTS: PR00078; G3PDHRCNASE.
CC DR PROSITE: PS00071; GAPDH: 1.
CC KW Glycolysis; Oxidoreductase; NAD.
CC FT INIT_MET 0 BY SIMILARITY.
CC FT BINDING 149 149 GLYCERALDEHYDE 3-PHOSPHATE.
CC FT ACT_SITE 176 176 ACTIVATES THIOL GROUP DURING CATALYSIS.
CC SQ SEQUENCE 332 AA; 35527 MW; A8D35010BC3039C7 CRC64;

Query Match 88.4%; Score 38; DB 1; Length 332;
Best Local Similarity 87.5%; Pred. No. 6.3;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDFGFRIG 8
Db 5 VNGFGRIG 12

RESULT 7
G3P_CRIGR STANDARD; PRT; 332 AA.
AC P17244;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12) (GAPDH).
GN GAPD.
OS Cricetus griseus (Chinese hamster).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
CC OC Cricetus.
CC RN NCBI_TaxID=10029;
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Lung;
CC RA MEDLINE-90272420; PubMed-2349105;
CC RT Vincent S., Fort P.;
CC RL "Nucleotide sequence of hamster glyceraldehyde-3-phosphate
CC RL dehydrogenase mRNA.";
CC RL Nucleic Acids Res. 18:3054-3054(1990).
CC CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC CC + NAD(+) -> 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC CC -1- SUBUNIT: HOMOTETRAMER.
CC CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC CC DEHYDROGENASE FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL: X52123; CAA36368.1; -

```

DR PIR: s10221: DEHYG.
 DR HSSP: P00354: 3GPD.
 DR InterPro: IPR000173: -
 DR Pfam: PF00044: gpch. 1.
 DR PRINTS: PR00078: G3PDHGNASE.
 DR PROSITE: PS00071: GAPDH. 1.
 KW Glycolysis: Oxidoreductase; NAD.
 FT INIT_MET 0
 FT BINDING 149 149 GLYCERALDEHYDE 3-PHOSPHATE.
 FT ACT_SITE 176 176 ACTIVATES THIOL GROUP DURING CATALYSIS.
 SQ SEQUENCE 332 AA: 35617 MW: 8F970C395FD824B1 CRC64;

Query Match 88.4%; Score 38; DB 1; Length 332;
 Best Local Similarity 87.5%; Pred. No. 6.3;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDFGFRIG 8
 1:|||||
 DB 5 VNGFGRIG 12

RESULT 8
 G3P_MOUSE STANDARD: PRT: 332 AA.
 AC P16858:
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12) (GAPDH).
 GN GAPD.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91007274; PubMed=2145197;
 RA Sebat D.E., Broome H.E., Prysicowsky M.B.;
 RT "Glyceraldehyde-3-phosphate dehydrogenase mRNA is a major Interleukin
 2-induced transcript in a cloned T-helper lymphocyte.";
 RL Gene 91:185-191(1990).
 CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
 + NAD(+) -> 1,3-DIPHOSPHATEGLYCERATE + NADH.
 CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M32599; AAA37659.1; -
 DR EMBL: M32599; AAA37659.1; -
 DR PIR: J70553; DEMSC.
 DR HSSP: P00354; 3GPD.
 DR SWISS-2DPAGE: P16858; MOUSE.
 DR MCD: MGI:95640: Gapd.
 DR InterPro: IPR000173: -
 DR Pfam: PF00044: gpch. 1.
 DR PRINTS: PR00078: G3PDHGNASE.
 DR PROSITE: PS00071: GAPDH. 1.
 KW Glycolysis: Oxidoreductase; NAD.
 FT INIT_MET 0
 FT BINDING 149 149 GLYCERALDEHYDE 3-PHOSPHATE.
 FT ACT_SITE 176 176 ACTIVATES THIOL GROUP DURING CATALYSIS.
 SQ SEQUENCE 332 AA: 35679 MW: 53123085BACF65D CRC64;

Query Match 88.4%; Score 38; DB 1; Length 332;
 Best Local Similarity 87.5%; Pred. No. 6.3;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDFGFRIG 8
 1:|||||
 DB 5 VNGFGRIG 12

RESULT 9
 G3P_PIG STANDARD: PRT: 332 AA.
 AC P00355; Q29546; P79299; P79317; O18816;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12) (GAPDH).
 GN GAPD.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=68399311; PubMed=4299800;
 RA Harris J.I., Perham R.N.;
 RT "Glyceraldehyde 3-phosphate dehydrogenase from pig muscle.";
 RL Nature 219:1025-1028(1968).
 RN [2]
 RP REVISION TO 45.
 RA Harris J.I., Davidson B.E., Sajgo M., Noller H.F., Perham R.N.;
 RL (In) Shugar D. (eds.);
 RL Enzymes and isoenzymes: structure, properties and function, pp.1-15,
 RL Academic Press, London and New York (1970).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Behdjani R., Silversides D.W.;
 RT "Sus scrofa glyceraldehyde-3-phosphate dehydrogenase, genomic and cDNA
 sequences.";
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-22.
 RC TISSUE=Liver.
 RX MEDLINE=76087882; PubMed=1201027;
 RA Kulbe R.D., Jackson K.W., Tang J.;
 RT "Structural evidence for a liver-specific glyceraldehyde-3-phosphate
 dehydrogenase.";
 RL Biochem. Biophys. Res. Commun. 67:35-42(1975).
 RN [5]
 RP SEQUENCE OF 1-125 FROM N.A.
 RC TISSUE=Small Intestine;
 RA Winkler A.K., Fredholm M.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 14-300 FROM N.A.
 RA Foss D.L., Murtough M.P.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 298-332 FROM N.A.
 RC TISSUE=Skeletal muscle;
 RA Davoli R., Zambonelli P., Fontanesi L., Bigl D., Costosi E., Russo V.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE OF 173-322 FROM N.A.
 RA Yelich J.V., Pomp D., Geisert R.D.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
 + NAD(+) -> 1,3-DIPHOSPHATEGLYCERATE + NADH.
 CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE

DEHYDROGENASE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AF017079; CAA94053.1;
 DR EMBL: X94251; CAA63935.1;
 DR EMBL: U48832; AAB91804.1;
 DR EMBL: U82261; AAB40155.1;
 DR EMBL: 284063; CAB06323.1;
 DR PIR: A00367; DEPG3.
 DR PIR: B12055; B12055.
 DR HSSP: P00354; 3GPD.
 DR InterPro: IPR000173;
 DR Pfam: PF00044; gpdh. 1.
 DR PRINTS: PR00078; G3PDHGNASE.
 DR PROSITE: PS00071; GAPDH; 1.
 KW Glycolysis; Oxidoreductase; NAD.
 FT BINDING 149 149
 FT ACT_SITE 176 176
 FT CONFLICT 6 6
 FT CONFLICT 61 61
 FT CONFLICT 67 67
 FT CONFLICT 69 70
 FT CONFLICT 81 81
 FT CONFLICT 89 90
 FT CONFLICT 90 90
 FT CONFLICT 132 132
 FT CONFLICT 144 144
 FT CONFLICT 164 164
 FT CONFLICT 200 200
 FT CONFLICT 222 222
 FT CONFLICT 235 235
 FT CONFLICT 276 276
 FT CONFLICT 281 281
 FT CONFLICT 285 286
 FT CONFLICT 310 310
 FT CONFLICT 320 320
 SQ SEQUENCE 332 AA; 35705 MW; PE6A1638C1EB4A30 CRC64;
 Glyceralddehyde 3-phosphate.
 ACTIVATES THIOL GROUP DURING CATALYSIS.
 N -> D (IN REF. 1).
 N -> D (IN REF. 1).
 N -> D (IN REF. 1).
 N -> NP (IN REF. 3).
 N -> K (IN REF. 3).
 AT -> TA (IN REF. 1).
 T -> E (IN REF. 3).
 V -> I (IN REF. 3).
 H -> N (IN REF. 3).
 A -> L (IN REF. 3).
 A -> D (IN REF. 1).
 P -> A (IN REF. 3).
 D -> H (IN REF. 3).
 C -> S (IN REF. 3).
 SD -> DS (IN REF. 1).
 W -> S (IN REF. 3).
 R -> S (IN REF. 8).

Query Match 88.4%; Score 38; DB 1; Length 332;
 Best Local Similarity 87.5%; Pred. No. 6.3;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGGRIG 8
 1:|||||
 Db 5 VNGGRIG 12

RESULT 10

G3P_RABIT
 ID G3P_RABIT STANDARD; PRT; 332 AA.
 AC P46406;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12) (GAPDH).
 GN GAPD.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=96011658; PubMed=7590291;

RA Applequist S.E., Keyna U., Calvin M.R., Beck-Engeser G.B., Raman C.,
 RA Jaek H.-M.;
 RT "Sequence of the rabbit glyceraldehyde-3-phosphate dehydrogenase-
 RT encoding cDNA.";
 RL Gene 163:325-326(1995).
 RN [2]

RX SEQUENCE OF 32-78 FROM N.A.
 RX MEDLINE=83167564; PubMed=6687628;
 RA Putney S.D., Herlitz W.C., Schimmel P.R.;
 RT "A new tropomyosin T and cDNA clones for 13 different muscle proteins,
 RT found by shotgun sequencing.";
 RL Nature 302:718-721(1983).
 CC -1 CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
 CC + NAD(+) = 1,3-DIPHOSPHATEGLYCERATE + NADH.
 CC -1 SUBUNIT: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
 CC -1 SUBUNIT: HOMOTETRAMER.
 CC -1 SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1 SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 CC DATABASE: NAME=worthington enzyme manual;
 CC WWW="http://www.worthington-biochem.com/manual/G/GAPD.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L23961; AAB85218.1;
 DR EMBL: V00884; CAA24253.1;
 DR HSSP: P00354; 3GPD.
 DR InterPro: IPR000173;
 DR Pfam: PF00044; gpdh. 1.
 DR PRINTS: PR00078; G3PDHGNASE.
 DR PROSITE: PS00071; GAPDH; 1.
 KW Glycolysis; Oxidoreductase; NAD.
 FT BINDING 149 149
 FT ACT_SITE 176 176
 FT CONFLICT 176 176
 SQ SEQUENCE 332 AA; 35688 MW; 1AB5051B1722A19A CRC64;
 BY SIMILARITY.
 Glyceralddehyde 3-phosphate.
 ACTIVATES THIOL GROUP DURING CATALYSIS.

Query Match 88.4%; Score 38; DB 1; Length 332;
 Best Local Similarity 87.5%; Pred. No. 6.3;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGGRIG 8
 1:|||||
 Db 5 VNGGRIG 12

RESULT 11

G3P_RAT
 ID G3P_RAT STANDARD; PRT; 332 AA.
 AC P04797; P09328;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12) (GAPDH) (38 KDA
 DE BFA-DEPENDENT ADP-RIBOSYLATION SUBSTRATE) (BARS-38).
 GN GAPD.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85215556; PubMed=2987824;
 RA Fort P., Marty L., Piechaczyk M., el Sabrouly S., Dani C.,
 RA Jeantaur P., Blanchard J.M.;
 RT "Various rat adult tissues express only one major mRNA species from

RT the glyceraldehyde-3-phosphate-dehydrogenase multigenic family";
 RL Nucleic Acids Res. 13:1431-1442(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-85215629; PubMed-2987855;
 RT Tso J.Y., Sun X.-H., Rao T.-H., Reece K.S., Wu R.;
 RT "Isolation and characterization of rat and human glyceraldehyde-3-
 RT phosphate dehydrogenase cDNAs: genomic complexity and molecular
 RT evolution of the gene."; Nucleic Acids Res. 13:2485-2502(1985).
 RN [3]
 RP SEQUENCE OF 260-322 FROM N.A.
 RA MEDLINE-85014145; PubMed-6548307;
 RT Piechaczyk M., Blanchard J.M., Marty L., Dani C., Panblieres F.,
 RT el Sabrouzy S., Fort P., Jeanne P.,
 RT "Post-transcriptional regulation of glyceraldehyde-3-phosphate-
 RT dehydrogenase gene expression in rat tissues."; Nucleic Acids Res. 12:6951-6963(1984).
 RN [4]
 RP SEQUENCE OF 234-332 FROM N.A.
 RC TISSUE=Brain;
 RA MEDLINE-87224934; PubMed-3585333;
 RT Leung T.K.C., Hall C., Montiles C., Lim L.;
 RT "Trifluoperazine activates and releases latent ATP-generating enzymes
 RT associated with the synaptic plasma membrane."; J. Neurochem. 49:232-238(1987).
 RN [5]
 RP SEQUENCE OF 266-332 FROM N.A.
 RA MEDLINE-86023533; PubMed-2413848;
 RT Maehara Y., Fujiyoshi T., Takahashi K., Yamamoto M., Endo H.;
 RT "1.5 kb mRNA abundantly expressed in rat tumors encodes a 37
 RT kilodalton protein in vitro."; Biochem. Biophys. Res. Commun. 131:800-805(1985).
 RL Bloecher, B. *et al.* Res. Commun. 131:800-805(1985).
 CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
 CC + NAD(+) = 1,3-DIPHOSPHATEGLYCERATE + NADH.
 CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- PTM: ADP-RIBOSYLATED BY BREFFELDIN-A (BFA).
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X02331; CAA26150.1; -
 DR EMBL: M29341; AAA40814.1; -
 DR EMBL: M11561; AAA41795.1; -
 DR EMBL: M17701; AAA41193.1; -
 DR PIR: A23013; DEPTG.
 DR HSSP: P00354; 3GPD.
 DR InterPro: IPR000173; -
 DR Pfam: PF00044; gpdh; 1.
 DR PRINTS: PR00078; G3PDHGRGNSE.
 DR PROSITE: PS00071; GAPDH; 1.
 KW Glycolysis; Oxidoreductase; NAD; ADP-ribosylation.
 DR INIT_MET 0
 FT BINDING 149 149 GLYCERALDEHYDE 3-PHOSPHATE.
 FT ACT_SITE 176 176 ACTIVATES THIOL GROUP DURING CATALYSIS.
 FT CONFLICT 80 81 VK -> AN (IN REF. 2).
 FT CONFLICT 304 304 I -> F (IN REF. 2, 4 AND 5).
 SO SEQUENCE 332 AA; 35705 MW; 343DEC24271CCEFE CRC64;

QY 1 VDGFGRIG 8
 DB 5 VNGFGRIG 12
 RESULT 12
 G3P_ZYGR0 STANDARD; PRT; 333 AA.
 AC P08439;
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DE 15-JUL-1999 (Rel. 38, Last annotation update)
 OS ZYGOSACCHAROMYCES rouxii (Candida mogii).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Zygosaccharomycetes.
 OX NCBI_TaxID=4956;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Imura T., Utsu T., Toh-E A.;
 RT "Glyceraldehyde-3-phosphate dehydrogenase genes of Zygosaccharomycetes
 RT rouxii: the source of a promoter for a host-vector system for Z.
 RT rouxii."; Agric. Biol. Chem. 51:1641-1647(1987).
 RL -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
 CC + NAD(+) = 1,3-DIPHOSPHATEGLYCERATE + NADH.
 CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D00134; BAA0081.1; -
 DR PIR: S00152; DEK2GR.
 DR HSSP: P00357; 1GPD.
 DR InterPro: IPR000173; -
 DR Pfam: PF00044; gpdh; 1.
 DR PRINTS: PR00078; G3PDHGRGNSE.
 DR PROSITE: PS00071; GAPDH; 1.
 KW Glycolysis; Oxidoreductase; NAD.
 FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE.
 FT ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS.
 SO SEQUENCE 333 AA; 35601 MW; 48B166336DD3DBC2 CRC64;

Query Match 88.4%; Score 38; DB 1; Length 333;
 Best Local Similarity 87.5%; Pred. No. 6.3;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGFGRIG 8
 DB 6 VNGFGRIG 13
 RESULT 13
 G3P2_HUMAN STANDARD; PRT; 334 AA.
 AC P04406;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, LIVER (EC 1.2.1.12).
 GN GAPD.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89008430; PubMed-3170585;
 RA Ercolani L., Florence B., Denaro M., Alexander M.;
 RT "Isolation and complete sequence of a functional human
 RL glycereraldehyde-3-phosphate dehydrogenase gene.";
 RN J Biol. Chem. 263:15335-15341(1988).
 [2]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Liver;
 MEDLINE-85215629; PubMed-2987855;
 RA Tso J.Y., Sun X.-H., Kao T., Reece K.S., Wu R.;
 RT "Isolation and characterization of rat and human glyceraldehyde-3-
 RL phosphate dehydrogenase cDNAs: genomic complexity and molecular
 evolution of the gene.";
 RN Nucleic Acids Res. 13:2485-2502(1985).
 [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-85076585; PubMed-6096136;
 RA Hanauer A., Mandel J.L.;
 RT "The glyceraldehyde 3 phosphate dehydrogenase gene family: structure
 RL of a human cDNA and of an X chromosome linked pseudogene; amazing
 EMBO J. 3:2627-2633(1984).
 [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-85087928; PubMed-6096821;
 RA Arcari P., Martinelli R., Salvatore F.;
 RT "The complete sequence of a full length cDNA for human liver
 RL glyceraldehyde-3-phosphate dehydrogenase: evidence for multiple mRNA
 species.";
 RN Nucleic Acids Res. 12:9179-9189(1984).
 [5]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Placenta;
 MEDLINE-92020872; PubMed-1924305;
 RA Meyer-Slegler K., Mauro D.V., Seal G., Wurzer J., Dertel J.K.,
 RT "A human nuclear uracil DNA glycosylase is the 37-kDa subunit of
 RL glyceraldehyde-3-phosphate dehydrogenase.";
 RN Proc. Natl. Acad. Sci. U.S.A. 88:8460-8464(1991).
 [6]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Lung;
 MEDLINE-88026722; PubMed-3664468;
 RA Tokunaga K., Nakamura Y., Sakata K., Fujimori K., Ohkubo M.,
 RT "Enhanced expression of a glyceraldehyde-3-phosphate dehydrogenase
 RL gene in human lung cancers.";
 RN Cancer Res. 47:5616-5619(1987).
 [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87109159; PubMed-3027061;
 RA Allen R.W., Trach K.A., Hoch J.A.;
 RT "Identification of the 37-kDa protein displaying a variable
 RL interaction with the erythroid cell membrane as
 glyceraldehyde-3-phosphate dehydrogenase.";
 RN J Biol. Chem. 262:649-653(1987).
 [8]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Brain;
 MEDLINE-20403900; PubMed-10944468;
 RA Ye Z., Connor J.R.;
 RT "cDNA cloning by amplification of circularized first strand cDNAs
 RL reveals non-IRE-regulated iron-responsive mRNAs.";
 RN Biochem. Biophys. Res. Commun. 275:223-227(2000).
 [9]
 RP SEQUENCE FROM N.A.
 RX Zhou J., Yu W., Tang H., Mei G., Tsang Y.T.M., Bouck J., Gibbs R.A.,
 RA Marzolin J.F.;
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RN [10]
 RP SEQUENCE OF 219-225 AND 241-245.
 RC TISSUE=Heart;
 RX MEDLINE-96007936; PubMed-7498159;
 RA Kovalyov L.I., Shishkin S.S., Elimochkin A.S., Kovalyova M.A.,
 RT Ershova E.S., Egorov T.A., Musalyamov A.K.;
 RL "The major protein expression profile and two-dimensional protein
 RT database of human heart.";
 CC Electrophoresis 16:1160-1169(1995).
 CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
 CC + NAD(+) = 1,3-DIPHOSPHATEGLYCERATE + NADH.
 CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: M17851; AAA86283.1; -
 DR EMBL: X01677; CAA25833.1; -
 DR EMBL: J04038; AAA53191.1; -
 DR EMBL: X53778; CAA37794.1; -
 DR EMBL: M33197; AAA52518.1; -
 DR EMBL: J02642; AAA52406.1; -
 DR EMBL: AF261085; AAF99678.1; -
 DR EMBL: AY001133; AAG01996.1; -
 DR PIR: A00365; DEHUGL.
 DR PIR: B22939; B22939.
 DR PIR: A21939; A21939.
 DR PIR: A31988; A31988.
 DR PIR: A41297; A41297.
 DR PIR: A45924; A45924.
 DR HSP: P00354; 3CPD.
 DR SWISS-2DPAGE: P04406; HUMAN.
 DR Aarhu/Chent: 2DPAGE; 1206; NEPHGE.
 DR HSC-2DPAGE: P04406; HUMAN.
 DR MIM: 138400; -
 DR InterPro: IPR000173; -
 DR Pfam: PF00044; gpdh.1.
 DR PRINTS: PR00078; G3PDHGNASE.
 DR PROSITE: PS00071; GAPDH; 1.
 KW Glycolysis; Oxidoreductase; NAD; Multigene family.
 FT INIT_MET 0
 FT BINDING 151 151
 FT ACT_SITE 178 178
 FT CONFLICT 224 224
 FT SEQUENCE 334 AA; 35922 MW; 6CE6C4BAF54C377 CRC64;
 N -> D (1N REF. 4).

Query Match 88.4%; Score 38; DB 1; Length 334;
 Best Local Similarity 87.5%; Pred. No. 6.3;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDSGRIG 8
 Db 7 VNGFGRIG 14

RESULT 14
 G3P1_JACOR
 ID G3P1_JACOR STANDARD; PRT; 363 AA.
 AC P80534; 064418;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, MUSCLE (EC 1.2.1.12)

```

DE (GAPDH).
OS Jaculus orientalis.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Dipodidae; Dipodinae;
OC Jaculus.
OX NCBI_TaxID=48868;
RN [1]
RP SEQUENCE OF 1-54.
RC TISSUE-SKELETAL muscle;
RA Soukri A., Serrano A.;
RL Submitted (JAN-1996) to the SWISS-PROT data bank.
RN [2]
RP SEQUENCE OF 25-363 FROM N.A.
RC TISSUE-SKELETAL muscle;
RX MEDLINE-97128782; PubMed-8973322;
RA Soukri A., Valverde F., Hafid N., Elkebdaj M.S., Serrano A.;
RT "Occurrence of a differential expression of the glyceraldehyde-3-
RT phosphate dehydrogenase gene in muscle and liver from euthermic and
RT induced hibernating jerboa (Jaculus orientalis).";
RL Gene 181:139-145(1996).
CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC + NAD(+) -> 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X87226; CAA60678.1; -.
DR HSP: P00354; 3GPD.
DR InterPro: IPR00173; -.
DR Pfam: PF00044; gpdh; 1.
DR PRINTS: PR00078; G3PDHGNASE.
DR PROSITE: PS00071; GAPDH; 1.
KM GLYCOLYSIS; Oxidoreductase; NAD.
FT BINDING 180 180 GLYCERALDEHYDE 3-PHOSPHATE
FT (BY SIMILARITY).
FT ACT_SITE 207 207 ACTIVATES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
SQ SEQUENCE 363 AA; 39413 MW; 6CB549D9A1163C3F CRC64;

Query Match 88.4%; Score 38; DB 1; Length 363;
Best Local Similarity 87.5%; Pred. NO. 6.8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDFGFRIG 8
   1:|||||
Db 6 VNGFGRIG 13

RESULT 15
G3P1_ANASP
ID G3P1_ANASP STANDARD: PRT: 35 AA.
AC P80506;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE 1 (EC 1.2.1.12) (GAPDH 1)
DE (FRAGMENT).
GN GAP1.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]

```

```

RP SEQUENCE.
RA Valverde F., Serrano A.;
RL Submitted (NOV-1995) to the SWISS-PROT data bank.
CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC + NAD(+) -> 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR HSP: P00362; 2GDI.
DR InterPro: IPR00173; -.
DR PROSITE: PS00071; GAPDH; PARTIAL.
KW Glycolysis; Oxidoreductase; NAD; Multigene family.
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3788 MW; D6CBFA2CCD908F9 CRC64;

Query Match 86.0%; Score 37; DB 1; Length 35;
Best Local Similarity 75.0%; Pred. NO. 1.3;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDFGFRIG 8
   1:|||||
Db 7 VNGFGRIG 14

```

Search completed: June 13, 2001, 14:30:38
Job time: 528 sec

Wed Jun 13 15:00:52 2001

pct-us01-05825a-30.rsp

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:29:46 ; Search time 150.43 Seconds
(without alignments)
6.233 Million cell updates/sec

Title: PCT-US01-05825A-30

Perfect score:

Sequence: 1 VDFGFRIG 8

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP unclassified:*
13: SP vertebrate:*
14: SP virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	88.4	29	11 090YB1	090YB1 rattus sp.
2	38	88.4	75	6 097653	097653 equus caball
3	38	88.4	76	6 097652	097652 felis silve
4	38	88.4	83	11 P97617	P97617 rattus norv
5	38	88.4	106	6 09M205	09M205 ovis aries
6	38	88.4	333	6 09N2D6	09N2D6 canis famli
7	38	88.4	333	6 09N2D5	09N2D5 felis silve
8	38	88.4	333	11 090WU4	090WU4 rattus norv
9	38	88.4	376	10 09M7R7	09M7R7 odontella s
10	38	88.4	379	10 09M7R4	09M7R4 phaeodactyl
11	38	88.4	871	10 09M907	09M907 arabisdopsi
12	38	88.4	1107	10 09ZUL9	09ZUL9 arabisdopsi
13	37	86.0	32	5 09NUT5	09NUT5 leishmania
14	37	86.0	39	2 09R5J2	09R5J2 streptococ
15	37	86.0	49	5 P90514	P90514 acanthamoeb
16	37	86.0	69	5 01S558	01S558 trichomonas
17	37	86.0	69	5 01S559	01S559 trichomonas
18	37	86.0	70	3 09URG7	09URG7 kluyveromyc
19	37	86.0	87	10 041949	041949 arabisdopsi

20	37	86.0	105	3 09Y7H1	09Y7H1 piromyces s
21	37	86.0	113	10 09SPK8	09SPK8 lactuca sat
22	37	86.0	113	10 09SEC3	09SEC3 lactuca sat
23	37	86.0	165	11 003850	003850 rattus norv
24	37	86.0	211	5 09NA02	09NA02 daphnia mag
25	37	86.0	290	10 023139	023139 arabisdopsi
26	37	86.0	308	2 056094	056094 salmonella
27	37	86.0	308	2 003594	003594 salmonella
28	37	86.0	308	2 003595	003595 salmonella
29	37	86.0	308	2 003596	003596 salmonella
30	37	86.0	308	2 003597	003597 salmonella
31	37	86.0	310	2 09X6Z1	09X6Z1 burkholderi
32	37	86.0	315	2 09WML6	09WML6 synechococc
33	37	86.0	320	2 0915X6	0915X6 streptococ
34	37	86.0	330	2 09RUP1	09RUP1 deinococcus
35	37	86.0	331	2 09K0J8	09K0J8 vibrio chol
36	37	86.0	331	5 09V318	09V318 drosophila
37	37	86.0	332	2 092KT0	092KT0 helicobacte
38	37	86.0	332	2 09PM04	09PM04 campylobact
39	37	86.0	332	5 09VXM7	09VXM7 drosophila
40	37	86.0	332	5 09N9Y3	09N9Y3 daphnia pul
41	37	86.0	333	2 09X7H5	09X7H5 paracoccus
42	37	86.0	333	2 09WML7	09WML7 synechococc
43	37	86.0	333	2 09KT36	09KT36 vibrio chol
44	37	86.0	333	13 09PTW5	09PTW5 paratichthy
45	37	86.0	334	2 052631	052631 clostridium

ALIGNMENTS

RESULT 1	
090YB1	PRELIMINARY: PRT; 29 AA.
AC 090YB1;	
DT 01-MAY-2000 (TREMBLrel. 13, Created)	
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)	
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)	
DE NITRIC OXIDE-ENHANCED ADP-RIBOSYLATED PROTEIN (FRAGMENT).	
OS Rattus sp.	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX NCBI_TaxID=10118;	
RN [1]	
RP SEQUENCE.	
RX MEDLINE=93028468; PubMed=1409644;	
RA Zhang J., Snyder S.H.;	
RT "Nitric oxide stimulates auto-ADP-ribosylation of glyceraldehyde-3-	
RT phosphate dehydrogenase."	
RL Proc. Natl. Acad. Sci. U.S.A. 89:9382-9385(1992).	
DR INTERPRO: IPR000173; -.	
DR PFWA; PF000044; gpdh; 1.	
SQ SEQUENCE 29 AA; 3109 MW; 5133EB197FC96EEA CRC64;	
Query Match	Score 38; DB 11; Length 29;
Best Local Similarity	87.5%; Pred. No. 1;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
OY 1 VDFGFRIG 8	
DB 1:	
5 VDFGFRIG 12	
RESULT 2	
097653	PRELIMINARY: PRT; 75 AA.
AC 097653;	
DT 01-MAY-1999 (TREMBLrel. 10, Created)	
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)	
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)	
DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (FRAGMENT).	
OS Equus caballus (Horse).	

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Leutenegger C.M., Mislin C., Akenis M., von Rechenberg B., Auer J.;
 RT "Partial characterization of equine glyceraldehyde-3-phosphate
 RL dehydrogenase (GAPDH).";
 DR EMBL; AF097178; AAC72841.1; to the EMBL/GenBank/DBJ databases.
 DR HSSP; P00354; 3GPD.
 DR INTERPRO; IPR000173;
 DR PFAM; PF00044; gpdh; 1.
 FT NON_TER 1 75
 FT SEQUENCE 75 AA; 8300 MW; C943E24C88E83AF8 CRC64;

Query Match 88.4%; Score 38; DB 6; Length 75;
 Best Local Similarity 87.5%; Pred. No. 3.1;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGFGRIG 8
 Db 6 VNGFGRIG 13

RESULT 3
 ID 097652 PRELIMINARY; PRT; 76 AA.
 AC 097652;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (FRAGMENT).
 OS Fells silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Fells.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Leutenegger C.M., Mislin C., Lutz H.;
 RT "Partial characterization of feline glyceraldehyde-3-phosphate
 RL dehydrogenase (GAPDH).";
 DR EMBL; AF097177; AAC72840.1; to the EMBL/GenBank/DBJ databases.
 DR HSSP; P00354; 3GPD.
 DR INTERPRO; IPR000173;
 DR PFAM; PF00044; gpdh; 1.
 FT NON_TER 1 76
 FT SEQUENCE 76 AA; 8334 MW; D3F3E09146FF7936 CRC64;

Query Match 88.4%; Score 38; DB 6; Length 76;
 Best Local Similarity 87.5%; Pred. No. 3.1;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGFGRIG 8
 Db 7 VNGFGRIG 14

RESULT 4
 ID P97617 PRELIMINARY; PRT; 83 AA.
 AC P97617;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (FRAGMENT).
 GN GAPDH.
 OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-WISTAR-KYOTO, TISSUE-AORTIC SMOOTH MUSCLE;
 RA Adams L.A., Werny I., Schwartz S.M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U75401; AAB19105.1; to the EMBL/GenBank/DBJ databases.
 DR HSSP; P17721; 1HDG.
 DR INTERPRO; IPR000173;
 DR PFAM; PF00044; gpdh; 1.
 FT NON_TER 83 83
 FT SEQUENCE 83 AA; 9240 MW; 73CF62461B4DCB61 CRC64;

Query Match 88.4%; Score 38; DB 11; Length 83;
 Best Local Similarity 87.5%; Pred. No. 3.5;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGFGRIG 8
 Db 6 VNGFGRIG 13

RESULT 5
 ID 09M205 PRELIMINARY; PRT; 106 AA.
 AC 09M205;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE GAPDH (FRAGMENT).
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Herrmann L.M., Davys W.C., Li H., Wardrup J., Sy M.-S., Gambetti P.,
 RA Knowles D.P.;
 RT "Detection of PrPC in peripheral blood mononuclear cells of normal
 RT sheep.";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF272837; AAF91410.1; to the EMBL/GenBank/DBJ databases.
 DR HSSP; P17721; 1HDG.
 FT NON_TER 1 106
 FT SEQUENCE 106 AA; 11612 MW; B22E924D623D550A CRC64;

Query Match 88.4%; Score 38; DB 6; Length 106;
 Best Local Similarity 87.5%; Pred. No. 4.6;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGFGRIG 8
 Db 5 VNGFGRIG 12

RESULT 6
 ID 09N2D6 PRELIMINARY; PRT; 333 AA.
 AC 09N2D6;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE.
 GN GAPDH.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ADRENAL GLAND;
 RA Kitemura H., Adachi K., Kido N., Hagiya T., Minase K., Yasui H.,
 RA Yano E., Ohta Y., Tabu K., Mae J., Kanehira K., Ohashi A.;
 RT "Canine glyceraldehyde-3-phosphate dehydrogenase (GAPDH), complete
 RT CDS."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB038240; BAA90817.1; -
 SQ SEQUENCE 333 AA; 35861 MW; E08BAC39FA2AE7EE CRC64;

Query Match 88.4%; Score 38; DB 6; Length 333;
 Best Local Similarity 87.5%; Pred. No. 17;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDFGFRIG 8
 Db 6 VNGFGRIG 13

RESULT 7
 Q9N2D5 PRELIMINARY; PRT; 333 AA.
 AC Q9N2D5:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE.
 GN GAPDH.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-WHITE ADIPOSE TISSUE;
 RA Kitemura H., Adachi K., Ohta Y., Kido N., Hagiya T., Yasui H.,
 RA Yano E., Minase K., Mae J., Tabu K., Kanehira K., Ohashi A.;
 RT "Feline glyceraldehyde-3-phosphate dehydrogenase (GAPDH), complete
 RT CDS."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB038241; BAA90818.1; -
 SQ SEQUENCE 333 AA; 35813 MW; 20501C401BBE7906 CRC64;

Query Match 88.4%; Score 38; DB 6; Length 333;
 Best Local Similarity 87.5%; Pred. No. 17;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDFGFRIG 8
 Db 6 VNGFGRIG 13

RESULT 8
 Q9QWU4 PRELIMINARY; PRT; 333 AA.
 AC Q9QWU4:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12).
 GN GAPDH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-BRAIN;
 RA Zheng J., Ramirez V.D.;
 RT "Cloning of a rat brain glyceraldehyde-3-phosphate dehydrogenase cDNA

RT by a rapid PCR-based method."
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF106860; AAD08929.1; -
 DR HSSP: P00354; 3GPD.
 DR INTERPRO: IPR000173; -
 DR PFAM: PF000044; gpdh; 1.
 DR PRINTS: PR00078; G3PDHGRGNASE.
 DR PROSITE: PS00071; GAPDH; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 333 AA; 35844 MW; 3185EADAA544180 CRC64;

Query Match 88.4%; Score 38; DB 11; Length 333;
 Best Local Similarity 87.5%; Pred. No. 17;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDFGFRIG 8
 Db 6 VNGFGRIG 13

RESULT 9
 Q9M7R7 PRELIMINARY; PRT; 376 AA.
 AC Q9M7R7:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE PRECURSOR.
 GN GAPC1.
 OS Odontella sinensis.
 OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
 OC Bidulophycidae; Eupodiscaceae; Odontella.
 OX NCBI_TaxID=2839;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lland M.-F., Lichte C., Apt K., Martin W., Cerff R.;
 RT "Compartment-Specific Isoforms of TPI and GAPDH are Imported Into
 RT Diatom Mitochondria as a Fusion Protein: Evidence in Favor of a
 RT Mitochondrial Origin of the Eukaryotic Glycolytic Pathway."
 RL Mol. Biol. Evol. 17:213-223(2000).
 DR EMBL: AF063800; AAF34326.1; -
 KW TRANSIT peptide.
 FT CHAIN 1 37 POTENTIAL.
 FT TRANSIT 38 376 GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE.
 SQ SEQUENCE 376 AA; 39931 MW; 73ADEF84E98FC75 CRC64;

Query Match 88.4%; Score 38; DB 10; Length 376;
 Best Local Similarity 87.5%; Pred. No. 20;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDFGFRIG 8
 Db 43 VNGFGRIG 50

RESULT 10
 Q9M7R4 PRELIMINARY; PRT; 379 AA.
 AC Q9M7R4:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE PRECURSOR.
 GN GAPC1.
 OS Phaeodactylum tricornutum.
 OC Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 OC Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 OX NCBI_TaxID=2850;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lland M.-F., Lichte C., Apt K., Martin W., Cerff R.;

RT "Compartment-Specific Isoforms of TPI and GAPDH are Imported into
RT Dlatom Mitochondria as a Fusion Protein: Evidence in Favor of a
RT Mitochondrial Origin of the Eukaryotic Glycolytic Pathway."
RL Mol. Biol. Evol. 17:213-223(2000).
DR EMBL; AF065803; AAF34329.1; -
RW Transit peptide.
FT TRANSIT
FT CHAIN 1 40 POTENTIAL.
FT SEQUENCE 379 AA: 40182 MW: 913931BIE66EE302 CRC64;

Query Match
Best Local Similarity 88.4%; Score 38; DB 10; Length 379;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGFGRIG 8
Db 46 VNGFGKIG 53

RESULT 11
O9M907 PRELIMINARY; PRT; 871 AA.
AC O9M907;
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DE F17A9.7 PROTEIN.
GN F17A9.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
ON NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Uteback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F17A9 genomic sequence."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC016827; AAF26996.1; -
SQ SEQUENCE 871 AA: 98338 MW: 08CE8F9297AD927 CRC64;

Query Match
Best Local Similarity 88.4%; Score 38; DB 10; Length 871;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGFGRIG 8
Db 664 IDGFGKIG 671

RESULT 12
O9ZU19 PRELIMINARY; PRT; 1107 AA.
AC O9ZU19;
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DE F504.8 PROTEIN.
GN F504.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
ON NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,

RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC F504 genomic sequence."
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005936; AAC97219.1; -
DR INTERPRO: IPR000504; -
DR INTERPRO: IPR002885; -
DR PFAM: PF01535; PPR: 16
DR PROSITE: PS00030; RNP_1; UNKNOWN_1.
SQ SEQUENCE 1107 AA: 124150 MW: 0C6D874CB23D5269 CRC64;

Query Match
Best Local Similarity 88.4%; Score 38; DB 10; Length 1107;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGFGRIG 8
Db 172 IDGFGKIG 179

RESULT 13
O9NUT5 PRELIMINARY; PRT; 32 AA.
AC O9NUT5;
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DE PUTATIVE GLYCOSOMAL GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE
DE (FRAGMENT).
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
ON NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RA Ch'ang L.Y., Lee T.;
RT "Trapping ORF-containing genomic fragments."
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF110677; AAF66094.1; -
FT NON_TER 32
SQ SEQUENCE 32 AA: 3406 MW: C9039F2CA020093A CRC64;

Query Match
Best Local Similarity 86.0%; Score 37; DB 5; Length 32;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGFGRIG 8
Db 8 INGFGRIG 15

RESULT 14
O9RSJ2 PRELIMINARY; PRT; 39 AA.
AC O9RSJ2;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DE SURFACE DEHYDROGENASE (FRAGMENT).
OS Streptococcus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae.
ON NCBI_TaxID=1301;
RN [1]
RP SEQUENCE.
RX MEDLINE=92364544; PubMed=1500854;
RA Pancholl V., Fischetti V.A.;
RT "A major surface protein on group A streptococci is a
RT glyceraldehyde-3-phosphate-dehydrogenase with multiple binding
RT activity."
RL J. Exp. Med. 176:415-426(1992).
DR HSSP: P00361. 1CER.
DR INTERPRO: IPR000173; -
DR PFAM: PF00044; gpdh; 1.

SQ SEQUENCE 39 AA; 4284 MW; 911ED3AF6B50F66F CRC64;

Query Match

Best Local Similarity 86.0%; Score 37; DB 2; Length 39;
Best Local Similarity 75.0%; Pred. No. 2.3;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDFGRIG 8
:::|||||

DB 6 INGFRIG 13

RESULT 15

P90514 PRELIMINARY; PRT; 49 AA.
AC P90514;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12) (GLYCERALDEHYDE
3-PHOSPHATE DEHYDROGENASE (PHOSPHORYLATING)) (TRIOSEPHOSPHATE
DEHYDROGENASE) (GAPDH) (FRAGMENT).
OS Acanthamoeba castellanii (Amoeba).
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
OX NCBI_TaxID-5755;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu P., Zot H.G.;
RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + PHOSPHATE +
CC NAD(+) = 3-PHOSPHO-D- GLYCEROYL PHOSPHATE + NADH.
DR EMBL: U85500; ABA1646.1; -.
DR HSSP: Q27890; 1GYP.
DR INTERPRO: IPR00173; -.
DR PFM: PF00044; gpdh; 1.
KW Oxidoreductase.
FT NON_TER 49
SQ SEQUENCE 49 AA; 5218 MW; 1EEF56C24855EA6B CRC64;

Query Match

Best Local Similarity 86.0%; Score 37; DB 5; Length 49;
Best Local Similarity 75.0%; Pred. No. 3;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDFGRIG 8
:::|||||

DB 8 INGFRIG 15

Search completed: June 13, 2001, 14:29:46
Job time: 547 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:27:08 ; Search time 78.71 Seconds
(without alignments)
1.953 Million cell updates/sec

Title: PCT-US01-05825A-30

Perfect score: 43

Sequence: 1 VDGRGRIG 8

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents, AA:*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	97.7	22	1	US-08-678-444-1
2	38	88.4	12	2	US-08-464-073-33
3	38	88.4	18	1	US-08-553-110-1
4	38	88.4	18	3	US-08-946-026-50
5	38	88.4	268	2	US-08-997-080-186
6	38	88.4	268	2	US-08-997-362-186
7	38	88.4	268	4	US-09-095-855-186
8	38	88.4	340	4	US-09-095-855-205
9	37	86.0	27	6	5245013-20
10	37	86.0	333	4	US-08-961-083-54
11	37	86.0	334	6	5290690-10
12	37	86.0	334	6	5290690-11
13	37	86.0	335	6	5290690-9
14	37	86.0	335	6	5290690-9
15	37	86.0	336	1	US-07-928-462-2
16	37	86.0	336	1	US-08-273-247-2
17	32	74.4	22	1	US-08-678-444-2
18	32	74.4	257	2	US-08-637-759B-92
19	32	74.4	257	3	US-08-871-355A-92
20	32	74.4	450	3	US-09-015-296-1
21	32	74.4	3798	4	US-09-335-408-6
22	31	72.1	36	2	US-08-461-990B-12
23	31	72.1	229	4	US-09-248-528-17
24	31	72.1	512	3	US-09-015-296-3
25	30	69.8	229	4	US-09-248-528-11
26	30	69.8	413	1	US-08-295-675A-2
27	30	69.8	413	1	US-08-750-077-6

28	30	69.8	413	2	US-08-948-591-2	Sequence 2, Appl
29	30	69.8	415	1	US-08-464-523B-27	Sequence 27, Appl
30	29	67.4	29	1	US-08-355-888A-7	Sequence 7, Appl
31	29	67.4	29	2	US-08-588-190-2	Sequence 2, Appl
32	29	67.4	29	2	US-08-693-697-7	Sequence 7, Appl
33	29	67.4	29	2	US-08-588-526-2	Sequence 2, Appl
34	29	67.4	29	2	US-08-640-389A-2	Sequence 2, Appl
35	29	67.4	29	3	US-08-693-696-7	Sequence 7, Appl
36	29	67.4	264	2	US-07-857-224B-9	Sequence 9, Appl
37	29	67.4	347	3	US-08-749-816-3	Sequence 3, Appl
38	29	67.4	411	2	US-08-440-845D-3	Sequence 3, Appl
39	29	67.4	411	2	US-08-440-845D-4	Sequence 4, Appl
40	29	67.4	411	4	US-08-868-458-3	Sequence 3, Appl
41	29	67.4	411	4	US-08-868-458-4	Sequence 4, Appl
42	29	67.4	415	1	US-08-464-523B-28	Sequence 28, Appl
43	29	67.4	715	4	US-08-669-286-5	Sequence 5, Appl
44	29	67.4	715	4	US-08-669-286-7	Sequence 7, Appl
45	29	67.4	715	4	US-09-469-253-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-678-444-1
Sequence 1, Application US/08678444
Patent No. 5679349
GENERAL INFORMATION:
APPLICANT: Scheifinger, Curtis C.
ATTORNEY/AGENT INFORMATION:
TITLE OF INVENTION: VACCINE DESIGN AND PRODUCTION
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company/1104
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: United States
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/678,444
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/147,765
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gaylo, Paul J.
REGISTRATION NUMBER: 36,808
REFERENCE/DOCKET NUMBER: X-8597
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317) 276-0756
TELEFAX: (317) 276-3861
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-678-444-1

Query Match 97.7%; Score 42; DB 1; Length 22;
Best local Similarity 87.5%; Pred. No. 0.05;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGRGRIG 8

Db 6 IDGFGRRIG 13

RESULT 2

US-08-464-073-33

Sequence 33, Application US/08464073

Patent No. 5864028

GENERAL INFORMATION:

APPLICANT: Sloud, Mouldy

TITLE OF INVENTION: TNF-ALPHA RIBOZYMES, TNF-ALPHA RIBOZYME BINDING

TITLE OF INVENTION: PROTEIN AND DEGRADATION RESISTANT MRNA DERIVATIVES

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESSES:

ADDRESS: Cooper and Dunham, LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/464,073

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/971,058

FILING DATE: 03-NOV-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/AU/1993/00567

FILING DATE: 03-NOV-1993

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-278-0400

TELEFAX: 212-391-0525

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-464-073-33

Query Match 88.4%; Score 38; DB 2; Length 12;
Best Local Similarity 87.5%; Pred. No. 0.15;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGFGRRIG 8
Db 5 VNGFGRRIG 12

RESULT 3

US-08-553-110-1

Sequence 1, Application US/08553110

Patent No. 5723301

GENERAL INFORMATION:

APPLICANT: Burke, James R.

APPLICANT: Vance, Jeffrey M.

APPLICANT: Enghild, Jan

APPLICANT: Strittmatter, Warren J.

TITLE OF INVENTION: Therapeutics for Diseases Associated

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESSES:

ADDRESS: Kenneth D. Sibley

STREET: P.O. Drawer 34009

CITY: Charlotte

STATE: No. 5723301th Carolina

COUNTRY: United States of America

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/553,110

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Sibley, Kenneth D.

REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5405-117

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-553-110-1

Query Match

Best Local Similarity

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGFGRRIG 8
Db 5 VNGFGRRIG 12

RESULT 4

US-08-946-026-50

Sequence 50, Application US/08946026

Patent No. 6034218

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Dillon, David C.

APPLICANT: Twardzik, Daniel R.

APPLICANT: Mitcham, Jennifer L.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESSES:

ADDRESS: SEED AND BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/946,026

FILING DATE: 07-OCT-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.424C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4500

TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-946-026-50

Query Match 88.4%; Score 38; DB 3; Length 18;
Best Local Similarity 87.5%; Pred. No. 0.23;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDFGGRIG 8
1:|||||
Db 7 VNGFGRIG 14

RESULT 5
US-08-997-080-186
Sequence 186, Application US/08997080
Patent No. 5968524
GENERAL INFORMATION:
APPLICANT: WATSON, JAMES D.
APPLICANT: TAN, PAUL L.J.
TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,080
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0563
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 186:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-080-186

Query Match 88.4%; Score 38; DB 2; Length 268;
Best Local Similarity 87.5%; Pred. No. 4;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDFGGRIG 8
1:|||||
Db 7 VNGFGRIG 14

RESULT 6
US-08-997-362-186

Sequence 186, Application US/08997362
Patent No. 5985287
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Hiyama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
APPLICANT: Scott, Linda
TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,362
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
FILING DATE: June 12, 1997
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
FILING DATE: August 29, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0563
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 186:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-362-186

Query Match 88.4%; Score 38; DB 2; Length 268;
Best Local Similarity 87.5%; Pred. No. 4;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDFGGRIG 8
1:|||||
Db 7 VNGFGRIG 14

RESULT 7
US-09-095-855-186
Sequence 186, Application US/09095855
Patent No. 6160093
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 208

CORRESPONDENCE ADDRESS:

ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet

REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:

INFORMATION FOR SEQ ID NO: 186:

SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-095-855-186

Query Match 88.4%; Score 38; DB 4; Length 268;
Best Local Similarity 87.5%; Pred. No. 4;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGFRIG 8
1:|||||||
DB 7 VNGFRIG 14

RESULT 8

US-09-095-855-205
Sequence 205, Application US/09095855
Patent No. 6160093

GENERAL INFORMATION:

APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/095,855
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet

REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:

INFORMATION FOR SEQ ID NO: 205:

SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-095-855-205

Query Match 88.4%; Score 38; DB 4; Length 340;
Best Local Similarity 87.5%; Pred. No. 5.2;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGFRIG 8
1:|||||||
DB 7 VNGFRIG 14

RESULT 9

5245013-20
Patent No. 5245013

APPLICANT: Ulevitch, Richard; Tobias, Peter
TITLE OF INVENTION: ACUTE PHASE PROTEIN MODULATING ENDOGENOUS
ACTIVITY OF LIPOPOLYSACCHARIDES, ASSAY METHODS AND POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/367,454
FILING DATE: 01-JUN-1989

APPLICATION NUMBER: 6,710
FILING DATE: 30-DEC-1986

APPLICATION NUMBER: 728,833
FILING DATE: 30-APR-1985

SEQ ID NO: 20;
LENGTH: 27

5245013-20

Query Match 86.0%; Score 37; DB 6; Length 27;
Best Local Similarity 75.0%; Pred. No. 0.55;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGFRIG 8
1:|||||||
DB 4 VNGFRIG 11

RESULT 10

US-08-961-083-54
Sequence 54, Application US/08961083
Patent No. 6159469

GENERAL INFORMATION:

APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: P8340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-54

Query Match 86.0%; Score 37; DB 4; Length 333;
Best Local Similarity 75.0%; Pred. No. 7.9;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDFGRIG 8
: : : : :
Db 6 INGFRIG 13

RESULT 11
5290690-10
PATENT NO. 5290690
APPLICANT: MRABET, NADIR; LASTERS, IGNACE; STANSSENS, PATRICK
ADDRESSEE: GASTON; WODAK, SHOSHANA; QUAX, WILHELMUS J.
STREET: TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
STABILITY OF PROTEINS
NUMBER OF SEQUENCES: 22
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/398.706
FILING DATE: 25-AUG-1989
SEQ ID NO: 10:
LENGTH: 334
5290690-10

Query Match 86.0%; Score 37; DB 6; Length 334;
Best Local Similarity 75.0%; Pred. No. 7.9;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDFGRIG 8
: : : : :
Db 6 INGFRIG 13

RESULT 12
5290690-11

PATENT NO. 5290690
APPLICANT: MRABET, NADIR; LASTERS, IGNACE; STANSSENS, PATRICK
ADDRESSEE: GASTON; WODAK, SHOSHANA; QUAX, WILHELMUS J.
STREET: TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
STABILITY OF PROTEINS
NUMBER OF SEQUENCES: 22
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/398.706
FILING DATE: 25-AUG-1989
SEQ ID NO: 11:
LENGTH: 334
5290690-11

Query Match 86.0%; Score 37; DB 6; Length 334;
Best Local Similarity 75.0%; Pred. No. 7.9;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDFGRIG 8
: : : : :
Db 6 INGFRIG 13

RESULT 13
US-08-903-800A-6
Sequence 6, Application US/08903800A
PATENT NO. 5935789
GENERAL INFORMATION:
APPLICANT: RHEE, Sang-Ki
APPLICANT: CHOI, Eun-Sung
APPLICANT: KIM, Chul-Ho
APPLICANT: SOHN, Jung-Hoon
APPLICANT: KANG, Hyun-Ah
APPLICANT: KIM, Hwa-Young
TITLE OF INVENTION: AUTONOMOUSLY REPLICATING SEQUENCES,
TITLE OF INVENTION: GAPDH GENE AND PROMOTER DERIVED FROM
TITLE OF INVENTION: HANSENIA POLYMORPH, EXPRESSION
TITLE OF INVENTION: VECTORS CONTAINING SAME AND METHOD FOR
TITLE OF INVENTION: THE SELECTION OF TRANSFORMANTS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: RHEE, Sang-Ki
STREET: Keukdong Villa Ka-101, Kwangjang-dong,
STREET: Kwangjin-gu
CITY: Seoul
STATE: Seoul
COUNTRY: Republic of Korea
ZIP: 143-210
ADDRESSEE: CHOI, Eun-Sung
STREET: Dasol Apt. 102-507, #395-3, Kung-dong,
STREET: Yuseong-gu
CITY: Daejeon
STATE: Daejeon
COUNTRY: Republic of Korea
ZIP: 305-335
ADDRESSEE: KIM, Chul-Ho
STREET: Shindonga Apt. 2-1101, Yongjeon-dong, Dong-gu
CITY: Daejeon
STATE: Daejeon
COUNTRY: Republic of Korea
ZIP: 300-200
ADDRESSEE: SOHN, Jung-Hoon
STREET: Nuri Apt. 103-506, Wolpyung-dong, Seo-gu
CITY: Daejeon
STATE: Daejeon
COUNTRY: Republic of Korea
ZIP: 302-280
ADDRESSEE: KANG, Hyun-Ah
STREET: Kyungseong Keunmaul Apt. 125-1501, Kalma-dong,
CITY: Daejeon
STATE: Daejeon
COUNTRY: Republic of Korea

pct-us01-05825a-30.ra1

Page 6

```

RESULT 15
US-07-928-462-2
; Sequence 2, Application US/07928462
; Patent No. 5328996
;
; GENERAL INFORMATION:
; APPLICANT: Boyle, Michael D.P.
; APPLICANT: Lottenberg, Richard
; APPLICANT: Broder, Christopher C.
; APPLICANT: von Marling, Gregory O.
; TITLE OF INVENTION: Bacterial Plasmid Receptors as
; TITLE OF INVENTION: Fibrinolytic Agents
; NUMBER OF SEQUENCES: 2
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Sallwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Review #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,462
; FILING DATE: 19920810
; CLASSIFICATION: 424
;
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/524,411
; FILING DATE: 16-MAY-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/330,849
; FILING DATE: 29-MAR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Sallwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: US/56S-13.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; US-07-928-462-2

```


This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:25:49 ; Search time 150.28 Seconds

(Without alignments)
8.368 Million cell updates/sec

Title: PCT-US01-05825A-31

Sequence: 1 XGLVLCVNGFGRIGRIGRLVYI 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0401.*

1: /SIDS6/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDS6/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	91	87.5	22	Y79132 Human foetal brain
2	66	63.5	124	Y29539 Human lung tumour
3	66	63.5	124	B44446 Human lung tumour
4	66	63.5	127	Y29540 Human lung tumour
5	66	63.5	127	B44447 Human lung tumour
6	66	63.5	335	Y07036 Breast cancer asso
7	66	63.5	20	Y05368 Human HCMV inducid
8	64	61.5	136	G54490 Zea mays protein f
9	63	60.6	21	G44842 Arabidopsis thalia
10	63	60.6	175	G28442 Arabidopsis thalia
11	63	60.6	337	Y53268 Lentinus edodes gl

12	63	60.6	338	18	W22503	Ubiqlitin-ribosoma
13	62	59.6	18	18	W27314	Human prostate ass
14	62	59.6	18	19	W45521	Glycerolaldehyde-3-p
15	62	59.6	18	20	Y02457	Fragment of glycer
16	62	59.6	39	14	R38788	Streptococcal surf
17	62	59.6	70	14	R40212	N-terminal of a pl
18	62	59.6	160	18	W14732	C. hirsutus GPD.
19	62	59.6	383	19	W55089	Streptococcus pneu
20	62	59.6	736	15	R56486	Plasmin receptor.
21	62	59.6	336	22	Y85681	Streptococcal plas
22	62	59.6	337	13	R22097	Glycerolaldehyde-3-p
23	62	59.6	338	15	R65965	T. niveum GAPDH.
24	62	59.6	369	11	R05738	glycerolaldehyde-3-p
25	61.5	59.1	93	21	G34289	Arabidopsis thalia
26	61.5	59.1	149	21	G35523	Arabidopsis thalia
27	61.5	59.1	174	21	G35522	Arabidopsis thalia
28	61.5	59.1	338	21	G17086	Arabidopsis thalia
29	61.5	59.1	364	21	G25110	Arabidopsis thalia
30	61.5	59.1	364	21	G17085	Arabidopsis thalia
31	61	58.7	334	21	Y55844	P. citrifil glycer
32	61	58.7	337	12	R12995	GAP-DH. Aspergill
33	61	58.7	420	21	G18019	Arabidopsis thalia
34	61	58.7	420	21	G51763	Arabidopsis thalia
35	61	58.7	422	21	G28827	Arabidopsis thalia
36	61	58.7	422	21	G32115	Arabidopsis thalia
37	61	58.7	438	21	G51345	Arabidopsis thalia
38	60	57.7	31	19	W53450	Arabidopsis thalia
39	59	56.7	107	21	G19017	Blifdibacterium ge
40	59	56.7	133	21	G33938	Zea mays protein f
41	59	56.7	135	21	G22593	Arabidopsis thalia
42	59	56.7	171	21	G22592	Zea mays protein f
43	59	56.7	268	20	Y14921	Zea mays protein f
44	59	56.7	333	21	B19020	Amino acid sequenc
45	59	56.7	333	21	B19021	Amino acid sequenc

ALIGNMENTS

RESULT 1	
ID	Y79132 standard; Peptide: 22 AA.
Y79132	
AC	Y79132:
XX	
DT	05-JUN-2000 (first entry)
XX	
DE	Human foetal brain zonulin N-terminal sequence.
XX	
KM	Zonulin; antagonist; zonula occludens toxin receptor;
KW	human; blood-brain barrier; antiinflammatory;
KM	gastrointestinal inflammation; therapy.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	Misc-difference 1
FT	Location/Qualifiers
PN	WO200007609-A1.
XX	
PD	17-FEB-2000.
XX	
PF	28-JUL-1999; 99WO-US16683.
XX	
PR	03-AUG-1998; 98US-0127815.
XX	
PA	(UYMA-) UNIV MARYLAND BALTIMORE.
XX	
PI	Fasano A;
XX	
DR	WPI; 2000-205565/18.
XX	

Query Match	87.5%	Score 91;	DB 21;	Length 22;
Best Local Similarity	90.5%;	Pred. No. 8.1e-07;		
Matches 19;	Conservative	0;	Mismatches 2.	Total

RESULT	2
Y29539	
ID	Y29539
XX	standard; Protein; 124 AA

Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine; immunotherapy; detection; inhibition.

W09938973-A2

26-JAN-1999; .99WO-US01642.

22-DEC-1998; 98US-0219245.

28-JAN-1998; 98US-0015029.

18-MAR-1998; 98US-0040831.

23-JUL-1998; 98US-0122192.

(COR1-) CORIXA CORP.

Frudakis TN, Lodes MJ, Mohamath R, Reed SC.

WPI; 1999-479187/40.

[illegible]

newly cancer specific polynucleotides for inhibiting the development of lung cancer

Example 3; Page 114: 171no. English

The present invention describes lung tumour specific polynucleotides and tumour antigens. 207144 to 207246 and 208301 to 208325 represent specifically claimed polynucleotides, and Y29486 to Y29571 represent amino acid sequences from the present invention. The lung tumour

50 Sequence 124 AA;

Matches	13; Conservative	1; Mismatches
---------	------------------	---------------

RESULT	3
B44446	
ID	B44446 standard; Protein; 124 AA
XX	

05-FEB-2001 (first entry)

human lung tumour-specific antigen encoded by cDNA #57.

lung cancer; lung cancer; cytostatic; vaccine

Homo sapiens.

PN WO200060077-A2.

PD 12-OCT-2000

30-MAR-2000; 2000WO-US08560.

02-APR-1999; 99US-0285323.

30-DEC-1999; 99US-0476235.

XX

1000

SECRET

N-PSDB; C79104.

Novel lung tumor

of treating cancer, especially lung cancer -

English

sequence is given in a specification relating to compounds for therapy and diagnosis of lung cancer. Polymers of

The polypeptides are useful for inhibiting the growth of a lung tumour protein are disclosed.

may be used to inhibit the development of

especially lung cancer.

Sequence 124 AA:

Best Local Similarity	63.58; Score 66; DB 21; Length 124;
Best Local Similarity	81.28; Score 66; DB 21; Length 124;

Matches	13;	Conservative	1;	Mismatches	2;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

[illegible]

Db 2 gkvkvngvfgrlgrl 17

RESULT 4
ID Y29540 standard; Protein: 127 AA.
XX Y29540;
AC Y29540;
XX 13-OCT-1999 (first entry)
DT
XX Human lung tumour protein L86S-51 predicted amino acid sequence.
DE
XX Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
KW Immunotherapy; detection; inhibition.
XX
OS Homo sapiens.
XX
PN WO9938973-A2.
PD 05-AUG-1999.
XX
XX 26-JAN-1999; 99WO-US01642.
XX
XX 22-DEC-1998; 98US-0219245.
PR 28-JAN-1998; 98US-0015022.
PR 28-JAN-1998; 98US-0015025.
PR 18-MAR-1998; 98US-0040828.
PR 18-MAR-1998; 98US-0040831.
PR 23-JUL-1998; 98US-0122191.
PR 23-JUL-1998; 98US-0122192.
XX
XX (CORI-) CORIXA CORP.
XX
PI Frudakis TN, Lodes MJ, Mohamath R, Reed SG;
DR WPI: 1999-479187/40.
DR N-PSDB: 207224.
XX
XX Lung tumour specific polynucleotides for inhibiting the development
PT of lung cancer
XX
XX Example 3; Page 114-115; 171pp; English.
XX
XX The present invention describes lung tumour specific polynucleotides
CC and tumour antigens. 207144 to 207246 and 208301 to 208325 represent
CC specifically claimed polynucleotides, and 29486 to 29571 represent
CC amino acid sequences from the present invention. The lung tumour
CC specific polynucleotides and polypeptides can be used in pharmaceutical
CC compositions and vaccines to inhibit the development of lung cancer.
CC They can also be used to detect lung cancer in a patient. Probes and
CC antibodies derived from the lung tumour sequences are useful in
CC detection of lung cancer.
XX
XX Sequence 127 AA:

Query Match 63.5%; Score 66; DB 20; Length 127;
Best Local Similarity 81.2%; Pred. No. 0.018;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 GLVLGVNGFGRIGRI 17
DB 20 gkvkvngvfgrlgrl 35

RESULT 5
ID B44447 standard; Protein: 127 AA.
XX B44447;
XX 05-FEB-2001 (first entry)

XX DE Human lung tumour-specific antigen encoded by cDNA #58.
XX
XX Lung tumour protein; lung cancer; cytostatic; vaccine.
XX
OS Homo sapiens.
XX
XX WO200060077-A2.
XX
XX 12-OCT-2000.
XX
XX 30-MAR-2000; 2000MO-US08560.
XX
XX 02-APR-1999; 99US-0285323.
PR 09-AUG-1999; 99US-0370838.
PR 30-DEC-1999; 99US-0476235.
PR 03-MAR-2000; 2000US-0518809.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Lodes MJ, Mohamath R, Secrist H;
PI WPI: 2000-638466/61.
XX
XX N-PSDB: C79105.
XX
XX Novel lung tumor polypeptides and polynucleotides, useful for
PT detecting, monitoring or treating cancer, especially lung cancer -
XX
XX Claim 1; Page 132; 243pp; English.
XX
XX The present sequence is given in a specification relating to compounds
CC for therapy and diagnosis of lung cancer. Polypeptides comprising at
CC least an immunogenic part of a lung tumour protein are disclosed.
CC The polypeptides are useful for inhibiting the development of cancer,
CC especially lung cancer. Samples of T cells expressing the polypeptides
CC may be used to inhibit the development of cancer. The polypeptides are
CC also useful for detecting and monitoring the progression of cancer,
CC especially lung cancer.
XX
XX Sequence 127 AA:

Query Match 63.5%; Score 66; DB 21; Length 127;
Best Local Similarity 81.2%; Pred. No. 0.018;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 GLVLGVNGFGRIGRI 17
DB 20 gkvkvngvfgrlgrl 35

RESULT 6
ID Y07036 standard; Protein: 335 AA.
XX Y07036;
XX
XX 02-JUL-1999 (first entry)
DT
XX Breast cancer associated antigen precursor sequence.
DE
XX Cancer associated antigen; diagnosis; research; treatment; human;
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW prostate cancer.
XX
XX Homo sapiens.
XX
XX WO9904265-A2.
XX
XX 28-JAN-1999.
XX
XX 15-JUL-1998; 98WO-US14679.
XX

PR 04-MAY-1999;	99US-0132484.
PR 05-MAY-1999;	99US-0132485.
PR 06-MAY-1999;	99US-0132486.
PR 07-MAY-1999;	99US-0132487.
PR 11-MAY-1999;	99US-0132863.
PR 14-MAY-1999;	99US-0134256.
PR 14-MAY-1999;	99US-0134218.
PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134768.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 16-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140693.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 01-JUL-1999;	99US-0142154.
PR 02-JUL-1999;	99US-0142055.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143542.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.

PR 22-JUL-1999;	99US-0145085.
PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 23-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 26-JUL-1999;	99US-0145224.
PR 27-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 28-JUL-1999;	99US-0145919.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 06-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0158293.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.

Oy	2	GLVLVGNGFGRIGRI	17
	1	: : : : : : : : :	
Dd	2	gxikigingfgfgrigr	17

G44842

ID	standard; Protein; 130 AA
XX	

DT 18-OCT-2000 (first entry)
XX

Lea mayas protein fragment SEQ ID NO: 56219

hybridisation assay; genetic mapping; gene expression control; termination sequence; corn.

Zea mays subsp. mays.

PN EP1033405-A2.

06-SEP-2000 PD

25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

09-MAR-1999; 99US-0123548.

25-MAR-1999; 99US-0126264.

01-APR-1999; 99US-0127462.

08-APR-1999; 99US-0128714.

19-APR-1999; 99US-0130077.

23-APR-1999; 99US-0130510.

28-APR-1999; 99US-0131449

30-APR-1999; 99US-0132407.

05-MAY-1999; 99US-0132485

06-MAY-1999; 99US-0132487

11-MAY-1999; 99US-01342863.

3905-0134218.

PR	14-MAY-1999;	9905-0134219
PR	14-MAY-1999;	9905-0134221
PR	14-MAY-1999;	9905-0134370
PR	18-MAY-1999;	9905-0134768
PR	19-MAY-1999;	9905-0134941
PR	20-MAY-1999;	9905-0135124
PR	21-MAY-1999;	9905-0135153
PR	24-MAY-1999;	9905-0135629
PR	25-MAY-1999;	9905-0136021
PR	27-MAY-1999;	9905-0136192
PR	01-JUN-1999;	9905-0136782
PR	03-JUN-1999;	9905-0137222
PR	04-JUN-1999;	9905-0137528
PR	07-JUN-1999;	9905-0137602
PR	08-JUN-1999;	9905-0137724
PR	10-JUN-1999;	9905-0138064
PR	10-JUN-1999;	9905-0138540
PR	14-JUN-1999;	9905-0138847
PR	16-JUN-1999;	9905-0139119
PR	16-JUN-1999;	9905-0139452
PR	17-JUN-1999;	9905-0139453
PR	17-JUN-1999;	9905-0139454
PR	18-JUN-1999;	9905-0139455
PR	18-JUN-1999;	9905-0139456
PR	18-JUN-1999;	9905-0139457
PR	18-JUN-1999;	9905-0139458
PR	18-JUN-1999;	9905-0139459
PR	18-JUN-1999;	9905-0139460
PR	18-JUN-1999;	9905-0139461
PR	18-JUN-1999;	9905-0139462
PR	18-JUN-1999;	9905-0139750
PR	18-JUN-1999;	9905-0139763
PR	21-JUN-1999;	9905-0139817
PR	22-JUN-1999;	9905-0139859
PR	23-JUN-1999;	9905-0140033
PR	24-JUN-1999;	9905-0140354
PR	26-JUN-1999;	9905-0140695
PR	28-JUN-1999;	9905-0140823
PR	30-JUN-1999;	9905-0140991
PR	01-JUL-1999;	9905-0141287
PR	01-JUL-1999;	9905-0141484
PR	02-JUL-1999;	9905-0142154
PR	06-JUL-1999;	9905-0142905
PR	08-JUL-1999;	9905-0142950
PR	09-JUL-1999;	9905-0142803
PR	12-JUL-1999;	9905-0142920
PR	13-JUL-1999;	9905-0142977
PR	14-JUL-1999;	9905-0143342
PR	15-JUL-1999;	9905-0143624
PR	16-JUL-1999;	9905-0144005
PR	16-JUL-1999;	9905-0144085
PR	19-JUL-1999;	9905-0144086
PR	19-JUL-1999;	9905-0144325
PR	19-JUL-1999;	9905-0144331
PR	19-JUL-1999;	9905-0144332
PR	19-JUL-1999;	9905-0144333
PR	19-JUL-1999;	9905-0144334
PR	19-JUL-1999;	9905-0144335
PR	20-JUL-1999;	9905-0144632
PR	20-JUL-1999;	9905-0144633
PR	21-JUL-1999;	9905-0144884
PR	21-JUL-1999;	9905-0144881
PR	21-JUL-1999;	9905-0145086
PR	22-JUL-1999;	9905-0145085
PR	22-JUL-1999;	9905-0145087
PR	22-JUL-1999;	9905-0145088
PR	23-JUL-1999;	9905-0145192
PR	23-JUL-1999;	9905-0145195
PR	23-JUL-1999;	9905-0145218
PR	23-JUL-1999;	9905-0145224

PR	26-JUL-1999;	9905-0145276
PR	27-JUL-1999;	9905-0145913
PR	27-JUL-1999;	9905-0145918
PR	28-JUL-1999;	9905-0145919
PR	28-JUL-1999;	9905-0145951
PR	02-AUG-1999;	9905-0146386
PR	02-AUG-1999;	9905-0146388
PR	02-AUG-1999;	9905-0146389
PR	03-AUG-1999;	9905-0147204
PR	04-AUG-1999;	9905-0147208
PR	04-AUG-1999;	9905-0147302
PR	05-AUG-1999;	9905-0147192
PR	05-AUG-1999;	9905-0147260
PR	06-AUG-1999;	9905-0147303
PR	06-AUG-1999;	9905-0147416
PR	09-AUG-1999;	9905-0147493
PR	09-AUG-1999;	9905-0147935
PR	10-AUG-1999;	9905-0148171
PR	11-AUG-1999;	9905-0148319
PR	12-AUG-1999;	9905-0148341
PR	13-AUG-1999;	9905-0148565
PR	13-AUG-1999;	9905-0148684
PR	16-AUG-1999;	9905-0148368
PR	17-AUG-1999;	9905-0148175
PR	18-AUG-1999;	9905-0148426
PR	20-AUG-1999;	9905-0148722
PR	20-AUG-1999;	9905-0149733
PR	20-AUG-1999;	9905-0149723
PR	23-AUG-1999;	9905-0149929
PR	23-AUG-1999;	9905-0149929
PR	25-AUG-1999;	9905-0149530
PR	25-AUG-1999;	9905-0150566
PR	26-AUG-1999;	9905-0150884
PR	27-AUG-1999;	9905-0151065
PR	27-AUG-1999;	9905-0151066
PR	27-AUG-1999;	9905-0151080
PR	30-AUG-1999;	9905-0151303
PR	31-AUG-1999;	9905-0151338
PR	01-SEP-1999;	9905-0151930
PR	07-SEP-1999;	9905-0152363
PR	10-SEP-1999;	9905-0153750
PR	13-SEP-1999;	9905-0153758
PR	15-SEP-1999;	9905-0154018
PR	16-SEP-1999;	9905-0154039
PR	20-SEP-1999;	9905-0154779
PR	22-SEP-1999;	9905-0155134
PR	23-SEP-1999;	9905-0155486
PR	24-SEP-1999;	9905-0155458
PR	28-SEP-1999;	9905-0156459
PR	29-SEP-1999;	9905-0156596
PR	04-OCT-1999;	9905-0157117
PR	05-OCT-1999;	9905-0157533
PR	06-OCT-1999;	9905-0157865
PR	07-OCT-1999;	9905-0158029
PR	08-OCT-1999;	9905-0158232
PR	12-OCT-1999;	9905-0158369
PR	13-OCT-1999;	9905-0158293
PR	13-OCT-1999;	9905-0159294
PR	13-OCT-1999;	9905-0159295
PR	14-OCT-1999;	9905-0158329
PR	14-OCT-1999;	9905-0159331
PR	14-OCT-1999;	9905-0159330
PR	14-OCT-1999;	9905-0159637
PR	14-OCT-1999;	9905-0159638
PR	18-OCT-1999;	9905-0159584
PR	21-OCT-1999;	9905-0160741
PR	21-OCT-1999;	9905-0160767
PR	21-OCT-1999;	9905-0160768
PR	21-OCT-1999;	9905-0160770
PR	21-OCT-1999;	9905-0160815
PR	22-OCT-1999;	9905-0160980
PR	22-OCT-1999;	9905-0160981
PR	22-OCT-1999;	9905-0160989

PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161350.
PR	26-OCT-1999;	99US-0161350.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161921.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match		
Best Local Similarity	60.6%;	Score 63; DB 21; Length 130;
Matches 10; Conservative	62.5%;	Pred. No. 0.05;
2 gkikigfgrtgrl 17	4; Mismatches	2; Indels 0; Gaps 0;


```

Oy      2 GLVGVNGFCRIGRI 17
        | : : ||||| :
Db       2 gkikigfgrtgrl 17

```


RESULT 10
G28442
ID G28442 standard; Protein; 175 AA.
XX AC G28442;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 33660.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KX termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128645.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139753.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 02-JUL-1999; 99US-0141842.
PR 06-JUL-1999; 99US-0142855.
PR 08-JUL-1999; 99US-0142890.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 21-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 10-AUG-1999; 99US-0147493.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149358.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0149932.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151388.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 23-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

XX W09733909-A2.
 XX 18-SEP-1997.
 XX 14-MAR-1997; 97WO-US04192.
 XX 11-APR-1996; 96US-0633840.
 XX 15-MAR-1996; 96US-0616745.
 XX (CORI-) CORIXA CORP.
 XX Dillion DC, Reed SG, Twardzik DR;
 XX WPI: 1997-470816/43.
 XX Immunogenic portions of prostate proteins - useful to develop
 XX products to detect, monitor, treat or inhibit development of
 XX prostate cancer
 XX Claim 12; Page 72; 84pp; English.
 XX 2 Peptides (W27314 and W27315) were obtained following screening of
 XX human prostate adenocarcinoma cell line LnCap, fgc
 XX (ATCC CRL 1740) with human prostaticitis sera. Comparison of the
 XX peptide sequences with database sequences revealed homology
 XX to polypeptides not previously associated with human prostate.
 XX These were glyceraldehyde-3-phosphate dehydrogenase (see W27316),
 XX alpha-human fructose biphosphate aldolase (see W27317),
 XX calreticulin (see W27318 and W27319), malate dehydrogenase (see
 XX W27320) and cystic disease fluid protein (see W27312 and W27313).
 XX Immunogenic portions of these peptides can be used in a claimed
 XX pharmaceutical composition for the treatment of prostate cancer, in
 XX a claimed vaccine for treatment of prostate cancer, or used to
 XX raise a monoclonal or polyclonal antibody useful in claimed methods
 XX for detecting prostate cancer or monitoring the progress of
 XX prostate cancer. 17 HPA polypeptides (see W2312-23 and W27303-07)
 XX and 14 other immunogenic polypeptides (W27308-21) are provided.
 XX Sequence 18 AA:

Query Match 59.6%; Score 62; DB 18; Length 18;
 Best Local Similarity 85.7%; Pred. NO. 0.009;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VLGVNGFGFRIGRI 17
 Db 3 VKVYNGFGFRIGRI 16

RESULT 14
 W45521
 ID W45521 standard; peptide; 18 AA.

AC W45521;
 DT 27-MAY-1998 (first entry)

DE Glyceraldehyde-3-phosphate dehydrogenase N-terminal sequence.

KW Glyceraldehyde-3-phosphate dehydrogenase; GAPDH; polyglutamine;
 KW CAG repeat domain; neurodegenerative disorder; inhibition;
 KW Huntington's disease; schizophrenia; psychiatric disorder.

XX Homo sapiens.

OS US5723301-A.

XX 03-MAR-1998.

PF 03-NOV-1995; 95US-0553110.

PR 03-NOV-1995; 95US-0553110.
 XX (UYDU-) UNITV DUKE.
 XX Burke JR, Enghild J, Strittmatter WJ, Vance JM;
 XX WPI: 1998-178475/16.
 XX Screening assay for inhibitors of GAPDH binding to polyglutamine -
 XX for use in treating neuro-degenerative diseases or psychiatric
 XX disorders
 XX Example 1; Column 12; 19pp; English.
 XX This is the N-terminal sequence of glyceraldehyde-3-phosphate
 XX dehydrogenase (GAPDH). The invention relates to a new screening assay
 XX for compounds capable of inhibiting binding of GAPDH to polyglutamine.
 XX The assay comprises providing an aqueous solution containing a test
 XX compound, polyglutamine-containing molecules and GAPDH and determining
 XX if binding of GAPDH to the polyglutamine-containing molecules is reduced
 XX in the presence of the test compound. Inhibitors identified by the above
 XX assay could be used to treat neurodegenerative diseases or psychiatric
 XX disorders caused by a gene containing an extended CAG domain, e.g.
 XX Huntington's disease or schizophrenia.
 XX Sequence 18 AA:

Query Match 59.6%; Score 62; DB 19; Length 18;
 Best Local Similarity 85.7%; Pred. NO. 0.009;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VLGVNGFGFRIGRI 17
 Db 1 VKVYNGFGFRIGRI 14

RESULT 15
 Y02457
 ID Y02457 standard; Peptide; 18 AA.

AC Y02457;

DT 13-JUL-1999 (first entry)

DE Fragment of glyceraldehyde-3-phosphate dehydrogenase.

KW N-terminal; prostate tumour cell; immunogenic; treatment; diagnosis;
 KW prostate cancer; glyceraldehyde-3-phosphate dehydrogenase.

OS Homo sapiens.

PN W09918210-A2.

XX 15-APR-1999.

PF 07-OCT-1998; 98WO-US21166.

PR 23-JUN-1998; 98US-0102679.

XX 07-OCT-1997; 97US-0946026.

PA (CORI-) CORIXA CORP.

PI Dillion DC, Mitcham JL, Reed SG, Twardzik DR;

XX WPI: 1999-277272/23.

XX New isolated prostate polypeptides useful for the treatment,
 XX diagnosis and monitoring of prostate cancer

PS Claim 12; Page 88; 106pp; English.

CC The present sequence represents the N-terminal of a human

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:23:13 ; Search time 87.97 Seconds
(without alignments)
17.187 Million cell updates/sec

Title: PCT-US01-05825A-31

Perfect score: 104

Sequence: 1 XGLVLGVNGFGRIGRIGRLVI 22

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR-67:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	63.5	335	1 DEHUG3	glyceraldhyde-3-p
2	65	62.5	338	2 S26976	glyceraldhyde-3-p
3	64	61.5	165	2 SI4477	glyceraldhyde-3-p
4	64	61.5	331	2 B48445	glyceraldhyde-3-p
5	64	61.5	333	1 DEHYG	glyceraldhyde-3-p
6	64	61.5	333	1 DEMSG	glyceraldhyde-3-p
7	64	61.5	333	1 DERTC	glyceraldhyde-3-p
8	64	61.5	333	1 DECHG3	glyceraldhyde-3-p
9	64	61.5	333	2 JC4309	glyceraldhyde-3-p
10	64	61.5	333	2 JC5370	glyceraldhyde-3-p
11	64	61.5	333	2 JN0678	glyceraldhyde-3-p
12	64	61.5	337	1 DEJCG	glyceraldhyde-3-p
13	64	61.5	337	1 DEYDCG	glyceraldhyde-3-p
14	64	61.5	337	2 S26946	glyceraldhyde-3-p
15	64	61.5	338	2 T47218	glyceraldhyde-3-p
16	63	60.6	22	2 B60475	glyceraldhyde-3-p
17	63	60.6	22	2 PNO163	glyceraldhyde-3-p
18	63	60.6	337	1 DEZMCG	glyceraldhyde-3-p
19	63	60.6	337	1 DEBHG	glyceraldhyde-3-p
20	63	60.6	337	2 A35080	glyceraldhyde-3-p
21	63	60.6	337	2 S26863	glyceraldhyde-3-p
22	63	60.6	337	2 S26973	glyceraldhyde-3-p
23	63	60.6	337	2 S40610	glyceraldhyde-3-p
24	63	60.6	338	1 DENDC	glyceraldhyde-3-p
25	63	60.6	340	2 T09663	glyceraldhyde-3-p
26	63	60.6	341	2 T08147	glyceraldhyde-3-p
27	62	59.6	21	2 B12055	glyceraldhyde-3-p
28	62	59.6	27	2 PA0083	glyceraldhyde-3-p
29	62	59.6	31	2 S13206	glyceraldhyde-3-p

30	62	59.6	34	2 A12055	glyceraldhyde-3-p
31	62	59.6	331	1 DEECG3	glyceraldhyde-3-p
32	62	59.6	331	1 DEUTC	glyceraldhyde-3-p
33	62	59.6	336	2 S71350	glyceraldhyde-3-p
34	62	59.6	336	2 A42963	glyceraldhyde-3-p
35	62	59.6	337	2 S26975	glyceraldhyde-3-p
36	62	59.6	337	2 S29814	glyceraldhyde-3-p
37	62	59.6	338	2 JN0452	glyceraldhyde-3-p
38	62	59.6	360	2 S38570	glyceraldhyde-3-p
39	61.5	59.1	338	2 JQ1287	glyceraldhyde-3-p
40	61	58.7	34	2 A60475	glyceraldhyde-3-p
41	61	58.7	333	2 JC6310	glyceraldhyde-3-p
42	61	58.7	335	2 T40292	glyceraldhyde-3-p
43	61	58.7	336	1 DEASG3	glyceraldhyde-3-p
44	61	58.7	336	2 T40235	glyceraldhyde-3-p
45	61	58.7	337	1 DEUSGM	glyceraldhyde-3-p

ALIGNMENTS

RESULT 1
DEHUG3
glyceraldhyde-3-phosphate dehydrogenase (EC 1.2.1.12) [validated] - human
N:Alternate names: triose phosphate dehydrogenase
N:Contains: uracil DNA glycosylase (EC 3.2.2.-), nuclear
C:Species: Homo sapiens (man)
C:Date: 02-Apr-1982 #sequence_revision 08-Dec-1994 #text_change 15-Sep-2000
C:Accession: A31988; A00365; A21939; I53309; B22939; A45924; I55258; A41297; S26758;
R:Ericolani, L.; Florence, B.; Denaro, M.; Alexander, M.
J. Biol. Chem. 263, 15335-15341, 1988
A:Title: Isolation and complete sequence of a functional human glyceraldehyde-3-phosph
A:Reference number: A31988; MUID:85008430
A:Accession: A31988
A:Molecule type: DNA
A:Residues: 1-335 <ERC>
A:Cross-references: GB:J04038; NID:G182980; PIDN:AA53191.1; PID:G182981
R:Rancari, P.; Martinelli, R.; Salvatore, F.
Nucleic Acids Res. 12, 9179-9189, 1984
A:Title: The glyceraldehyde 3 phosphate dehydrogenase gene family: structure of a hum
A:Reference number: A00365; MUID:85087928
A:Accession: A00365
A:Molecule type: mRNA
A:Residues: 1-224, 'D', 226-335 <ARC>
A:Cross-references: GB:X01677; GB:R03121; GB:M17851; GB:X01110; NID:G31644; PIDN:CAA2
R:Hanauer, A.; Mandel, J.L.
EMBO J. 3, 2627-2633, 1984
A:Title: The glyceraldehyde 3 phosphate dehydrogenase gene family: structure of a hum
A:Reference number: A21939; MUID:85076585
A:Accession: A21939
A:Molecule type: mRNA
A:Residues: 1-335 <HAN>
A:Cross-references: GB:X01677; GB:R03121; GB:M17851; GB:X01110; NID:G31644
R:Randi, C.; Piechaczyk, M.; Audigier, Y.; El Sabouty, S.; Cathala, G.; Marcy, L.; For
Eur. J. Biochem. 145, 299-304, 1984
A:Title: Characterization of the transcription products of glyceraldehyde 3-phosphate
A:Reference number: I53309; MUID:85053136
A:Accession: I53309
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 292-310, 'V', 312-335 <DAN>
A:Cross-references: GB:M28283; NID:G182978; PIDN:AA52519.1; PID:G182979
R:Tiso, J.Y.; Sun, X.H.; Kao, T.; Reece, K.S.; Wu, R.
Nucleic Acids Res. 13, 2485-2502, 1985
A:Title: Isolation and characterization of rat and human glyceraldehyde-3-phosphate d
A:Reference number: A93562; MUID:85215629
A:Accession: B22939
A:Molecule type: mRNA
A:Residues: 2-335 <TSO>
A:Cross-references: GB:M17851; NID:G182860; PIDN:AA6283.1; PID:G182861
R:Tokunaga, K.; Nakamura, Y.; Sakata, K.; Fujimori, K.; Okubo, M.; Sawada, K.; Sakiy
Cancer Res. 47, 5616-5619, 1987
A:Title: Enhanced expression of a glyceraldehyde-3-phosphate dehydrogenase gene in hu

A:Reference number: A45924; MUID:88026722
 A:Accession: A45924
 A:Molecule type: mRNA
 A:Residues: 1-335 <TOK>
 A:Cross-references: GB:M3197; NID:g182976; PIDN:AAA52518.1; PID:g182977
 J: Biol. Chem. 262, 649-653, 1987
 A:Title: Identification of the 37-kDa protein displaying a variable interaction with the
 A:Reference number: I55258; MUID:87109155
 A:Accession: I55258
 A:Status: translated from GB/EMBL/DBD
 A:Molecule type: mRNA
 A:Residues: 1-335 <ALL>
 A:Cross-references: GB:J02642; NID:g182862; PIDN:AAA52496.1; PID:g182863
 R: Meyer-Siegler, K.; Mauro, D.J.; Seal, G.; Wurzer, U.; derkiet, U.K.; Slrover, M.A.
 Proc. Natl. Acad. Sci. U.S.A. 88, 8460-8464, 1991
 A:Title: A human nuclear uracil DNA glycosylase is the 37-kDa subunit of glyceraledehyde-
 A:Reference number: A41297; MUID:92020872
 A:Accession: A41297
 A:Molecule type: mRNA
 A:Residues: 1-335 <MEY>
 A:Cross-references: GB:X53778
 R: Slrover, M.A.
 submitted to the EMBL Data Library, July 1990
 A:Reference number: S26758
 A:Accession: S26758
 A:Molecule type: mRNA
 A:Residues: 1-193; 'NCGYMA', 201, 203, 'SRTSSLPL', 213, 'L', 215-329, 333-335 <SIR>
 A:Cross-references: EMBL:X53778; NID:g35052; PIDN:CAA37794.1; PID:g35053
 R: Nowak, K.; Kuzek, M.; Ostropolska, L.; Malarska, A.; Wolny, M.; Branowski, T.
 Hoppe-Seyler's Z. Physiol. Chem. 356, 1181-1183, 1975
 A:Title: The covalent structure of glyceraledehyde-phosphate dehydrogenase from human mus
 A:Reference number: A12103; MUID:76067491
 A:Accession: A12103
 A:Molecule type: Protein
 A:Residues: 2-8, 'B', 10-27, 'B', 25-27, '46-47', '2', '49', 'B', 51-61, 67-69, 'B', 71, 'K', 'B', 82, 'A2',
 V', 171, '22', 174-176, 'SS', 179-180, 'AB', 186-188, 'B', 190-194, 'B', 199-202, 'A', 204-205, 'L', 20
 Z', 318-321, 323-334, 'SGGVK', <NOM>
 A:Note: Some of this partial sequence was assigned tentatively based on composition
 R: Nowak, K.; Wolny, M.; Banas, T.
 FEBS Lett. 134, 143-146, 1981
 A:Title: The complete amino acid sequence of human muscle glyceraledehyde 3-phosphate de
 A:Reference number: A00366; MUID:82073291
 A:Accession: A00366
 A:Molecule type: Protein
 A:Residues: 2-8, 'D', 10-40, 'H', 42-63, 'D', 65-69, 'D', 71, 'KA', 74-82, 'EN', 85-91, 'TA', 94-112, 'R',
 S10ud, M.; Jespersen, L.
 J. Mol. Biol. 237, 775-789, 1996
 A:Title: Enhancement of hammerhead ribozyme catalysis by glyceraledehyde-3-phosphate dehy
 A:Reference number: S66563; MUID:96194445
 A:Accession: S66563
 A:Molecule type: Protein
 A:Residues: 4-15 <STO>
 R: Mercer, W.D.; Winn, S.I.; Watson, H.C.
 J. Mol. Biol. 104, 277-283, 1976
 A:Title: Twinning in crystals of human skeletal muscle D-glyceraledehyde-3-phosphate dehy
 A:Reference number: A39927; MUID:76265083
 R: Watson, H.C.; Campbell, J.C.
 Submitted to the Brookhaven Protein Data Bank, June 1983
 A:Reference number: A50598; PDB:3GPD
 A:Contents: annotation; X-ray crystallography, 3.5 angstroms, residues 2-8, 'D', 10-40, 'H',
 168-189, 'S', 191-197, 'G', 199-202, 'A', 204-205, 'L', 207-224, 'D', 226-242, 'L', 244-263, 'E', 265-
 C:Comment: Although the gene for this protein is a member of a multigene family and ther
 only one gene appears to be expressed.
 C:Genetics:
 A:Gene: GDB:GAPD
 A:Cross-references: GDB:119249; OMIM:138400
 A:Map position: 12p13.31-12p13.1
 A:Introns: 10/2; 43/3; 79/2; 109/3; 148/2; 175/3; 313/2
 C:Complex: homotrimer; crystallizes as two symmetric dimers in which the partners have

C:Function: <CYT>
 A:Description: oxidoreductase; reversibly catalyzes the oxidative phosphorylation of
 A:Pathway: gluconeogenesis; glycolysis
 A:Note: tetrameric form; cytosol
 C:Function: <NMC>
 A:Description: DNA repair; uracil DNA glycosylase for base-excision in DNA repair
 A:Note: monomeric form; nuclear
 C:Superfamily: glyceraledehyde-3-phosphate dehydrogenase
 C:Keywords: cytosol; DNA repair; gluconeogenesis; glycolysis; glycosidase; homotetram
 E:2-35/Product: glyceraledehyde-3-phosphate dehydrogenase #status experimental <MAT>
 E:152/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:152/Active site: Cys #status experimental
 F:179/Active site: His #status predicted

Query Match 63.5%; Score 66; DB 1; Length 335;
 Best Local Similarity 81.2%; Pred. No. 0.029;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GLVAVGVNGFGRIGRI 17
 Db 2 GKAVGVNGFGRIGRI 17

RESULT 2

S26976

glyceraledehyde-3-phosphate dehydrogenase (EC 1.2.1.12) II - cultivated mushroom

C:Species: Agaricus bisporus (cultivated mushroom)

C:Date: 23-Apr-1993 #sequence_revision 23-Apr-1993 #text_change 11-Jun-1999

C:Accession: S26976

R:Harnsen, M.C.; Schuren, F.H.J.; Mookha, S.M.; van Zullen, C.M.; Punt, P.J.; Wessels

Curr. Genet. 22, 447-454, 1992

A:Title: Sequence analysis of the glyceraledehyde-3-phosphate dehydrogenase genes from

A:Reference number: S26973; MUID:93113715

A:Accession: S26976

A:Molecule type: DNA

A:Residues: 1-338 <HAR>

A:Cross-references: EMBL:M81728; NID:g166330; PIDN:AAA32634.1; PID:g166331

C:Genetics:

A:Gene: GPIII

C:Superfamily: glyceraledehyde-3-phosphate dehydrogenase

C:Keywords: gluconeogenesis; glycolysis; homotrimer; NAD; oxidoreductase

Query Match 62.5%; Score 65; DB 2; Length 338;
 Best Local Similarity 80.0%; Pred. No. 0.041;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 LVLVGVNGFGRIGRI 17
 Db 1 MVRVGVNGFGRIGRI 15

RESULT 3

S14477

glyceraledehyde-3-phosphate dehydrogenase homolog - rat (fragment)

N:Alternate names: glyceraledehyde-3-phosphate dehydrogenase-like protein

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Jun-1999

C:Accession: S14477; A24515

R:Krawetz, S.A.; Connor, W.; Dixon, G.H.

submitted to the EMBL Data Library, October 1990

A:Reference number: S14477

A:Accession: S14477

A:Molecule type: mRNA

A:Residues: 1-165 <KRA1>

A:Cross-references: EMBL:X54798; NID:g57571; PIDN:CAA38569.1; PID:g57572

R:Krawetz, S.A.; Connor, W.; Cannon, P.D.; Dixon, G.H.

DNA 5, 427-435, 1986

A:Title: A vector-primer-clonase-sequencer plasmid for the construction of cDNA librar

A:Reference number: A24915; MUID:87053179

A:Accession: A24915

A:Molecule type: mRNA

A:Residues: 1-52 <KRA2>
 A:Cross-references: GB:M4166; NID:9204191; PIDN:AAA1178.1; PID:9204192
 C:Genetics:
 A:Gene: G3PD
 C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase

Query Match 61.5%; Score 64; DB 2; Length 165;
 Best Local Similarity 72.2%; Pred. No. 0.027;
 Matches 13; Conservative 3; Mismatches 0; Indels 2; Gaps 1;

OY 6 VGVNGFGRIGR--IGRLV 21
 ||:|||||||:|:
 DB 4 VGINCGRIGRLVGRLL 21

RESULT 4
 B48445
 glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - Leishmania mexicana
 C:Species: Leishmania mexicana
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 11-Jun-1999
 C:Accession: B48445; S25142
 R:Hammett, V.; Blaauw, M.; Kohl, L.; Allert, S.; Oppendoes, F.R.; Michels, P.A.M.
 Mol. Biochem. Parasitol. 55, 115-126, 1992
 A:Title: Molecular analysis of the cytosolic and glycosomal glyceralddehyde-3-phosphate
 A:Reference number: A48445; MUID:93063042
 A:Accession: B48445
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-331 <HAN>
 A:Cross-references: EMBL:X65220; NID:99552; PIDN:CAA46323.1; PID:99553
 C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase
 C:Keywords: oxidoreductase

Query Match 61.5%; Score 64; DB 2; Length 331;
 Best Local Similarity 73.3%; Pred. No. 0.056;
 Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVLVGVNGFGRIGRI 17
 :|:|||||||:
 DB 1 MVKVGNGFGRIGRI 15

RESULT 5
 DEHYG
 glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - Chinese hamster
 C:Species: Crictetus griseus (Chinese hamster)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 11-Jun-1999
 C:Accession: S10221
 R:Vincent, S.; Fort, P.
 Nucleic Acids Res. 18, 3054, 1990
 A:Title: Nucleotide sequence of hamster glyceralddehyde-3-phosphate dehydrogenase mRNA.
 A:Reference number: S10221; MUID:9072420
 A:Accession: S10221
 A:Molecule type: mRNA
 A:Residues: 1-333 <VIN>
 A:Cross-references: EMBL:X52123; NID:949434; PIDN:CAA36368.1; PID:949435
 C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase
 C:Keywords: gluconogenesis; glycolysis; homotetramer; NAD; oxidoreductase
 F:3-33/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:150,177/Active site: Cys, His #status predicted

Query Match 61.5%; Score 64; DB 1; Length 333;
 Best Local Similarity 80.0%; Pred. No. 0.056;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVLVGVNGFGRIGRI 17
 :|:|||||||:
 DB 1 MVKVGNGFGRIGRI 15

RESULT 6
 DEMSG
 glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 11-Jun-1999
 C:Accession: J70553; S14160
 R:Sabath, D.E.; Broome, H.E.; Prysowsky, M.B.
 Gene 91, 185-191, 1990
 A:Title: Glyceralddehyde-3-phosphate dehydrogenase mRNA is a major interleukin 2-induc
 A:Reference number: J70553; MUID:91007274
 A:Accession: J70553
 A:Molecule type: mRNA
 A:Residues: 1-333 <SAB>
 A:Cross-references: GB:M32599; NID:9193423; PIDN:AAA37659.1; PID:9309243
 R:Fillipek, A.; Gerke, V.; Weber, K.; Kuznicki, J.
 Eur. J. Biochem. 195, 795-800, 1991
 A:Title: Characterization of the cell-cycle-regulated protein calyculin from Ehrlich
 A:Reference number: S14160; MUID:91153321
 A:Accession: S14160
 A:Molecule type: protein
 A:Residues: 2-20, 'FSCD', 25-26, 'D', 28-64, 'T', 66-70, 'F', 72-77, 'F', 79-80, 'VK', 83-86, 'D',
 C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase
 C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
 F:2-333/Product: glyceralddehyde-3-phosphate dehydrogenase #status experimental <MAT>
 F:3-33/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:150,177/Active site: Cys, His #status predicted

Query Match 61.5%; Score 64; DB 1; Length 333;
 Best Local Similarity 80.0%; Pred. No. 0.056;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVLVGVNGFGRIGRI 17
 :|:|||||||:
 DB 1 MVKVGNGFGRIGRI 15

RESULT 7
 DERTG
 glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - rat
 N:Alternate names: triosephosphate dehydrogenase
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Jun-1999
 C:Accession: A23013; A22939; A22887; A23280; A17135; JN0401; A50208; B17155
 R:Fort, P.; Marty, L.; Piechaczky, M.; El Sabouty, S.; Danl, C.; Jeanne, P.; Blanc
 Nucleic Acids Res. 13, 1431-1442, 1985
 A:Title: Various rat adult tissues express only one major mRNA species from the glyce
 A:Reference number: A23013; MUID:85215556
 A:Accession: A23013
 A:Molecule type: mRNA
 A:Residues: 1-333 <FOR>
 A:Cross-references: GB:X02231; NID:956187; PIDN:CAA26150.1; PID:956188
 R:Tso, J.Y.; Sun, X.H.; Kao, T.; Reece, K.S.; Wu, R.
 Nucleic Acids Res. 13, 2485-2502, 1985
 A:Title: Isolation and characterization of rat and human glyceralddehyde-3-phosphate d
 A:Reference number: A93562; MUID:85215629
 A:Accession: A22939
 A:Molecule type: mRNA
 A:Residues: 1-80, 'AN', 83-304, 'F', 306-333 <RNO>
 A:Cross-references: GB:M17701; NID:9204248; PIDN:AAA1193.1; PID:9204249
 R:Piechaczky, M.; Blanchard, J.M.; Marty, L.; Danl, C.; Panabieres, F.; El Sabouty, S
 Nucleic Acids Res. 12, 6951-6963, 1984
 A:Title: Post-transcriptional regulation of glyceralddehyde-3-phosphate-dehydrogenase
 A:Reference number: A22887; MUID:85014145
 A:Accession: A22887
 A:Molecule type: mRNA
 A:Residues: 261-323 <PIE>
 A:Cross-references: GB:X00972
 R:Maehara, Y.; Fujiyoshi, T.; Takahashi, K.; Yamamoto, M.; Endo, H.
 Biochem. Biophys. Res. Commun. 131, 800-805, 1985
 A:Title: 1.5 KB mRNA abundantly expressed in rat tumors encodes a 37 kilodalton prote
 A:Reference number: A23280; MUID:86025553
 A:Accession: A23280

A:Molecule type: mRNA
 A:Residues: 267-304, 'F', 306-333 <MAE>
 A:Cross-references: GB:M1561; NID:g205963; PIDN:AAA41795.1; PID:g205964
 R:Vospelnikova, N.D.; Safonova, M.I.; Shuvailova, E.R.; Baralova, L.A.; Kniazev, S.P.; N
 Blochm, J. 199, 757-765, 1981
 A:Title: Identification of an arginine residue important for catalytic activity in the F
 A:Reference number: A90313; MUID:82182080
 A:Accession: A17155
 A:Molecule type: protein
 A:Residues: 117-119, 'N', 121-122, 'T', 124-126, 'Z', 128, 'LF', 131-133, 'BRH', 138, 'SK', 294-297
 R:Baralova, L.A.; Zheleznova, A.O.; Belyanova, L.P.; Baralova, L.A.; Vospelnikova, N.D.; S
 Blochm, J. 199, 757-765, 1981
 A:Title: Peptide sequence containing the active site cysteine of D-glyceraldehyde-3-phos
 A:Reference number: JN0401
 A:Accession: JN0401
 A:Molecule type: protein
 A:Residues: 144-160 <BAT>
 A:Experimental source: skeletal muscle
 R:Leung, T.K.C.; Hall, C.; Monfries, C.; Lim, L.
 J. Neurochem 49, 232-238, 1987
 A:Title: Trifluoperazine activates and releases latent ATP-generating enzymes associated
 A:Reference number: A60208; MUID:8724934
 A:Accession: A60208
 A:Status: preliminary; translation not shown
 A:Molecule type: mRNA
 A:Residues: 235-304, 'F', 306-333 <LEU>
 A:Cross-references: GB:M2341; NID:g203141; PIDN:AAA40814.1; PID:g203142
 A:Experimental source: brain
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
 C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD: oxidoreductase
 F:3-33/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:150/Active site: Cys #status experimental
 F:177/Active site: His #status predicted

Query Match 61.5%; Score 64; DB 1; Length 333;
 Best Local Similarity 80.0%; Pred. No. 0.056;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Db 1 MWKGVNGFGRIGRI 17
 :| ||||| ||||| :
 1 MWKGVNGFGRIGRI 15

RESULT 8
 DECH3
 glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - chicken
 N:Alternate names: GAP dehydrogenase
 C:Species: Gallus gallus (chicken)
 C>Date: 03-Aug-1984 #sequence-revision 30-Sep-1989 #text-change 16-Jun-2000
 A:Accession: A00368; A32737; A22035; I50640
 R:Dugalczyk, A.; Haron, J.A.; Stone, E.M.; Dennison, O.E.; Rothblum, K.N.; Schwartz, R.
 Biochem Biophys Res Commun 118, 767-773, 1984
 A:Title: Cloning and sequencing of a deoxyribonucleic acid copy of glyceraldehyde-3-phos
 A:Reference number: A00368; MUID:83204755
 A:Accession: A00368
 A:Molecule type: mRNA
 A:Residues: 1-333 <DUG>
 A:Cross-references: GB:V00407; GB:J00849; NID:963402; PIDN:CAA23698.1; PID:91628381
 R:Panabieres, F.; Plechaczky, M.; Ratner, B.; Daul, C.; Fort, P.; Riad, S.; Marty, L.;
 Biochem Biophys Res Commun 118, 767-773, 1984
 A:Title: Complete nucleotide sequence of the messenger RNA coding for chicken muscle gly
 A:Reference number: A32737; MUID:84153854
 A:Accession: A32737
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-333 <PAN>
 A:Cross-references: GB:K01458; NID:g211800; PIDN:AAA48778.1; PID:g211801
 R:Stone, E.M.; Rothblum, K.N.; Alevy, M.C.; Kuo, T.M.; Schwartz, R.J.
 Proc Natl Acad Sci U S A 82, 1628-1632, 1985
 A:Title: Complete sequence of the chicken glyceraldehyde-3-phosphate dehydrogenase gene.
 A:Reference number: A22035; MUID:85166184
 A:Accession: A22035

A:Molecule type: DNA
 A:Residues: 1-293, 'H', 295-333 <STO>
 A:Cross-references: GB:M1213; NID:g211796; PIDN:AAA48774.1; PID:g211797
 A:Note: the authors translated the codon CAT for residue 294 as Asp
 R:Arnold, H.H.; Dondy, H.; Wlebaner, K.; Datta, K.; Siddiqui, M.A.O.
 J Biol Chem 257, 9872-9877, 1982
 A:Title: Cloning, partial sequencing, and expression of glyceraldehyde-3-phosphate de
 A:Reference number: I50231; MUID:82265644
 A:Accession: I50231
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 'G', 198-276, 'E', 278-333 <ARN>
 A:Cross-references: GB:J00848; NID:g211798; PIDN:AAA48777.1; PID:g211799
 A:Accession: I50640
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 'G', 198-276, 'E', 278-333 <ARN>
 A:Cross-references: EMBL:V00406; NID:g63400; PIDN:CAA23697.1; PID:g63401
 C:Genetics:
 A:Insertions: 8/2; 41/3; 77/3; 107/3; 146/2; 173/3; 224/2; 250/3; 278/3; 311/2
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
 C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD: oxidoreductase
 F:2-33/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:3-33/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:150/177/Active site: Cys, His #status predicted

Query Match 61.5%; Score 64; DB 1; Length 333;
 Best Local Similarity 80.0%; Pred. No. 0.056;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Db 1 MWKGVNGFGRIGRI 17
 :| ||||| ||||| :
 1 MWKGVNGFGRIGRI 15

RESULT 9
 JCA4309
 glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - rabbit
 N:Alternate names: triose phosphate dehydrogenase
 N:Contains: uracil DNA glycosylase (EC 3.2.2.-), nuclear
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C>Date: 16-Nov-1995 #sequence-revision 08-Feb-1996 #text-change 11-Jun-1999
 A:Accession: JCA4309; I46482
 R:Applequist, S.E.; Keyna, U.; Calvin, M.R.; Beck-Engeser, G.B.; Raman, C.; Jaek, H.
 Gene 163, 325-326, 1995
 A:Title: Sequence of the rabbit glyceraldehyde-3-phosphate dehydrogenase-encoding cDN
 A:Reference number: JCA4309; MUID:96011658
 A:Accession: JCA4309
 A:Molecule type: mRNA
 A:Residues: 1-333 <APP>
 A:Cross-references: GB:L23961; NID:g406106; PIDN:AAA85218.1; PID:g406107
 A:Experimental source: spleen
 R:Putney, S.D.; Herlihy, W.C.; Schimmel, P.
 Nature 302, 718-721, 1983
 A:Title: A new tropomyosin T and CDNA clones for 13 different muscle proteins, found by
 A:Reference number: I46471; MUID:83167564
 A:Accession: I46482
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 33-79 <PUT>
 A:Cross-references: EMBL:V00884; NID:91563; PIDN:CAA24253.1; PID:g929756
 C:Genetics:
 A:Gene: gapdh
 C:Complex: homotetramer; crystallizes as two symmetric dimers in which the partners h
 C:Function: monomer
 C:Function: <CYT>
 A:Description: oxidoreductase; reversibly catalyzes the oxidative phosphorylation of
 A:Note: tetrameric form; cytosol
 C:Function: <NUC>
 A:Description: DNA repair; uracil DNA glycosylase for base-excision in DNA repair
 A:Note: monomeric form; nuclear

C: Superfamily: glyceralddehyde-3-phosphate dehydrogenase
C: Keywords: cytosol; DNA repair; gluconeogenesis; glycolysis; glycosidase; homotetramer

Query Match	61.5%	Score 64;	DB 2;	Length 333;
Best Local Similarity	80.0%	Pred. No. 0.056;		
Matches 12;	Conservative 2;	Mismatches 1;	Indels 0;	Gaps 0;

```
QY      3 LVLVGNGFGRIGRI 17
        :| | | | | | | | |
Db      1 MVKVGNGFGRIGRI 15
```

RESULT 10

JCS5370
 glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12), euthermic tissue - desert jerboa
 N:Alternate names: GAPDH, triosephosphate dehydrogenase
 C:Species: *Jaculus orientalis* (desert jerboa)
 C:Date: 28-May-1997 #sequence_revision 12-Sep-1997 #text_change 20-Jun-2000
 C:Accession: JCS5370; PC4314; S62674
 R:Soukri, A.; Valverde, F.; Hafid, N.; Elkebbaj, M.S.; Serrano, A.
 Gene 181, 139-145, 1996
 A:Title: Occurrence of a differential expression of the glyceralddehyde-3-phosphate dehydrogenase in the liver of *Jaculus orientalis*
 I:Reference number: JCS5370; MUID:97128782

A:Accession: JCS370
A:Molecule type: mRNA
A:Residues: 1-33 <SOU1>
A:Cross-references: EMBL:X87226; NID:g1103589; PIDN:CA60678.1; PID:g1103590
A:Accession: PC4314
A:Molecule type: protein
A:Residues: 1-56 <SOU2>
A:Experimental source: skeletal muscle
R:Soukri, A.; Haïd, N.; Valverde, F.; Elkebbaj, M.S.; Serrano, A.
B:Title: Biophys. Acta 1292, 177-187, 1996
A:Title: Evidence for a posttranslational covalent modification of liver glyceraldehyde
A:Reference number: S62673; MUID:96135342
A:Accession: S62674
A:Molecule type: protein
A:Residues: 2-26 <SOU>
A:Genetics:
A:Gene: gapC
C:Complex: homotetramer
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: gluconeogenesis; glycolysis; hibernation; homotetramer; oxidoreductase
I:2-33/Products: glyceraldehyde-3-phosphate dehydrogenase #status experimental <NAT>

```
QY      3 LVLGVNGFGRI 17
        :| | | | | | | |
Db      1 MKVGVNGFGRI 15
```

RESULT 1.1

JN0678 glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - qual1
C:Species: *Colurnix coturnix* (qual1)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 11-Jun-1999
C:Accession: JN0678; 535726
R:Weiskirchen, R.; Stelmaster, G.; Hartl, M.; Bister, K.
Gene 128, 269-272, 1993
A:Title: Sequence and expression of glyceralddehyde-3-phosphate dehydrogenase-encoding
I:Reference number: JN0678; MUID:93292297

C;Keywords: gluconeogenesis; glycolysis; oxidoreductase

Query Match	61.58;	Score 64;	DB 2;	Length 333;
Best Local Similarity	80.08;	Pred. No. 0.056;		
Matches 12;	Conservative 2;	Mismatches 1;	Indels 0;	Gaps 0;

```
QY      3 LVLGVNGFGRIGRI 17
          :+ ++++++ :
Db      1 MKVGVNGFGRIGRI 15
```

RESULT 12

DEJGCG
glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - chestnut blight fungus
C:Species: Cryphonectria parasitica, Endothia parasitica (chestnut blight fungus)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Jun-1999
C:Accession: S11447
R:Choi, G.H.; Nuss, D.L.
Nucleic Acids Res. 18, 5566, 1990
A:Title: Nucleotide sequence of the glyceralddehyde-3-phosphate dehydrogenase gene from
A:Reference number: S11447; MUID:91016863
A:Accession: S11447
A:Molecule type: DNA
A:Residues: 1-337 <CHO>
A:Cross-references: EMBL:X53996; NID:g2638; PIDD:CAA37943.1; PID:g295939
C:Genetics:
A:Gene: gpd-1
A:Introns: 42/3; 271/1
C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase
C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F:1-34/Region: beta-alpha-beta NAD nucleotide-binding fold
F:151,178/Active site: Cys, His #status predicted

```
Query Match      61.5%;   Score 64;   DB 1;   Length 337;
Best Local Similarity 80.0%;   Pred. No. 0.057;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

Qy	3	LVLVGNFGGRIGRI	17
		:1111111111	
Db	2	VVKVGINGFGRIGRI	16

RESULT 13

deydcg
 glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - fungus (Curvularia lunata)
 C.Species: Curvularia lunata
 C.Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 11-Jun-1999
 C.Accession: S17981
 R.Ostlewaacz, H.D., Ridder, R.
 Curr. Genet. 20, 151-155, 1991
 A>Title: Genome analysis of imperfect fungi: electrophoretic karyotyping and character
 A.Reference number: S17981; MUID:92035062
 A.Accession: S17981
 A.Molecule type: DNA
 A.Residues: 1-337 <OS>
 A.Cross-References: EMBL:X58718; NID:92600; PIDN:CAA4154.1; PID:g2601
 C.Genetics:
 A:Gene: gpd
 A.Introns: 4/3; 21/2; 42/3; 271/1
 C.Superfamily: glyceralddehyde-3-phosphate dehydrogenase
 C.Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
 F:4-34/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:151,178/Active site: Cys, His #status predicted

Query Match	61.5%	Score 64	DB 1	Length 337
Best Local Similarity	80.0%	Pred. No. 0.057		
Matches 12	Conservative	2	Mismatches 1	Indels 0
			Gaps	0
OY	3	LVLGVNGEGRIGRI	17	
	:			

QY 3 LVLGVNGFGRIGRI 17
:| | | | | | | | | |

Db 2 VVKVGINGFGRI 16

RESULT 14

S26946

glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - fungus (*Cochliobolus heterostrophus*)
C:Species: *Cochliobolus heterostrophus*, *Bipolaris maydis*
A:Valley: strain C3

C:Date: 23-Apr-1993 #sequence_revision 23-Apr-1993 #text_change 11-Jun-1999
C:Accession: S26946; S22150

R:van Wert, S.L.; Yoder, O.C.

Curr. Genet. 22, 29-35, 1992

A:Title: Structure of the *Cochliobolus heterostrophus* glyceraldehyde-3-phosphate dehydrogenase
A:Reference number: S26946; MID:92306164

A:Accession: S26946

A:Molecule type: DNA

A:Residues: 1-337 <VAN>

A:Cross-references: EMBL:X63516; NID:92581; PID:CA45084.1; PID:92582

A>Note: the authors translated the codon GAC for residue 28 as Arg, AAG for residue 56 as
R:Yoder, O.

submitted to the EMBL Data Library, December 1991

A:Accession: S22150

A:Molecule type: DNA

A:Residues: 1-21, 'HRAQRRRCRRK', 35-36, 'HRAPLRS', 44-337 <YOD>

A:Cross-references: EMBL:X63516

C:Genetics:

A:Gene: gpd1

A:Introns: 4/3; 21/2; 42/3; 271/1

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase

Query Match

Best Local Similarity 61.5%; Score 64; DB 2; Length 337;

Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVLVGVNGFGRI 17

Db 2 VVKVGINGFGRI 16

RESULT 15

T47218

glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) [Imported] - *Neurospora crassa*

C:Species: *Neurospora crassa*

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 19-May-2000

C:Accession: T47218

R:Sahni, M.; Kinsey, J.A.

submitted to the EMBL Data Library, April 1996

A:Description: Isolation of the *Neurospora crassa* glyceraldehyde-3-phosphate dehydrogenase
A:Reference number: Z24398

A:Accession: T47218

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-338 <SAH>

A:Cross-references: EMBL:U56397; PID:AB00570.1

C:Genetics:

A:Gene: gpd-1

A:Map position: LG IIR

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: oxidoreductase

Query Match

Best Local Similarity 61.5%; Score 64; DB 2; Length 338;

Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Search completed: June 13, 2001, 14:23:13
Job time: 744 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:30:38 ; Search time 51.57 Seconds

(without alignments)
14.614 Million cell updates/sec

Title: PCT-US01-05825A-31

Perfect score: 104
Sequence: 1 XGLVGVGNFGRIGRIGRLVI 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	63.5	334	1 G3P2_HUMAN	P04406 homo sapien
2	65	62.5	338	1 G3P2_AGABI	P32636 agarticus bl
3	64	61.5	332	1 G3P3_ECOLI	P33898 escherichia
4	64	61.5	337	1 G3P3_COCHO	P29497 cochlidoct
5	64	61.5	337	1 G3P3_CRYPA	P19089 cryphonectr
6	64	61.5	337	1 G3P3_CURLU	P28844 curvularia
7	64	61.5	337	1 G3P3_LYOSH	Q92243 lyophyllum
8	64	61.5	338	1 G3P3_NEUCR	P54118 neurospora
9	64	61.5	363	1 G3P1_JACOR	P08735 zea mays orl
10	63	60.6	337	1 G3PC_MAIZE	P17878 mesembryant
11	63	60.6	337	1 G3PC_MESCR	Q42977 oryza sativ
12	63	60.6	337	1 G3PC_ORYSA	P09054 zea mays (m
13	63	60.6	337	1 G3PX_MAIZE	P26517 hordeum vul
14	63	60.6	337	1 G3PX_HORVU	P00584 clavigens p
15	63	60.6	337	1 G3P3_CLAPU	P32637 podospota a
16	63	60.6	337	1 G3P3_PODAN	P32638 schizophyll
17	63	60.6	337	1 G3P3_SCHCO	P26521 ranunculus
18	63	60.6	338	1 G3PC_RANAC	P13507 ranunculus
19	63	60.6	338	1 G3P3_PPARH	Q39766 ginkgo bil
20	63	60.6	340	1 G3PC_GINBI	P34924 pinus sylve
21	63	60.6	340	1 G3PC_PINSY	Q41595 taxus bacea
22	63	60.6	340	1 G3PC_TAXBA	P49644 chlamydomon
23	63	60.6	341	1 G3PC_CHLRE	P80447 jaculus orl
24	62	59.6	25	1 G3P2_JACOR	P10096 bos taurus
25	62	59.6	320	1 G3P3_BOVIN	P06977 escherichia
26	62	59.6	330	1 G3P1_ECOLI	P00158 leishmania
27	62	59.6	330	1 G3PC_LEIME	P10097 trypanosoma
28	62	59.6	330	1 G3PC_TRYBB	P00356 gallus gall
29	62	59.6	332	1 G3P3_CHICK	Q57479 columba liv
30	62	59.6	332	1 G3P3_GOLLI	P05025 coturnix co
31	62	59.6	332	1 G3P3_COTJA	P17244 cricetus
32	62	59.6	332	1 G3P3_CRIGR	P16858 mus musculu
33	62	59.6	332	1 G3P3_MOUSE	

ALIGNMENTS

RESULT	ID	STANDARD	PRT	334 AA
1	G3P2_HUMAN			
AC	P04406			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE, LIVER (EC 1.2.1.12).			
GN	GAPD.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89008430; PubMed=3170585;			
RA	Escollani L., Florence B., Denaro M., Alexander M.;			
RT	"Isolation and complete sequence of a functional human			
RT	glyceraldehyde-3-phosphate dehydrogenase gene.";			
RL	J. Biol. Chem. 263:15335-15341(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=85215629; PubMed=2987855;			
RA	Tso J.Y., Sun X.-H., Kao T., Reese K.S., Wu R.;			
RT	"Isolation and characterization of rat and human glyceraldehyde-3-			
RT	phosphate dehydrogenase cDNAs: genomic complexity and molecular			
RT	evolution of the gene.";			
RL	Nucleic Acids Res. 13:2485-2502(1985).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85076585; PubMed=6096136;			
RA	Hanauer A., Mandel J.L.;			
RT	"The glyceraldehyde 3 phosphate dehydrogenase gene family: structure			
RT	of a human cDNA and of an X chromosome linked pseudogene; amazing			
RT	complexity of the gene family in mouse.";			
RL	EMBO J. 3:2627-2633(1984).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85087928; PubMed=6096821;			
RA	Atcari P., Martinelli R., Salvatore F.;			
RT	"The complete sequence of a full length cDNA for human liver			
RT	glyceraldehyde-3-phosphate dehydrogenase: evidence for multiple mRNA			
RT	species.";			
RL	Nucleic Acids Res. 12:9179-9189(1984).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=92020872; PubMed=1924305;			
RA	Meyer-Siegler K., Mauro D.J., Seal G., Wurtzer J., Dertel J.K.,			
RT	Strover M.A.;			
RT	"A human nuclear uracil DNA glycosylase is the 37-kDa subunit of			
RT	glyceraldehyde-3-phosphate dehydrogenase.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:8460-8464(1991).			
RN	[6]			
RP	SEQUENCE FROM N.A.			

34	62	59.6	332	1 G3P3_PIG	P00355 sus scrofa
35	62	59.6	332	1 G3P3_RABIT	P46406 oryctolagus
36	62	59.6	332	1 G3P3_RAT	P04797 rattus norv
37	62	59.6	335	1 G3P3-STREQ	Q59906 streptococc
38	62	59.6	335	1 G3P3-STREP	P50467 streptococc
39	62	59.6	337	1 G3P1_AGABI	P32635 agarticus bl
40	62	59.6	337	1 G3P2_TRIKO	P17730 trichoderma
41	62	59.6	337	1 G3P3_COLDN	P54117 collettotric
42	62	59.6	337	1 G3P3_LACLA	P52987 lactococcus
43	62	59.6	338	1 G3P3_COLGL	P35143 collettotric
44	62	59.6	338	1 G3P3_ERYGR	P00640 erysiphe gr
45	62	59.6	338	1 G3P3_TRITHA	P87197 trichoderma

DR	Pfam:PF00044; gpddh_1.	
DR	PRINTS: PR00078; G3PDHGNASE.	
KR	PROSITE: PS00071; G3PDH_1.	
KM	GLYCOLYSIS: Oxidoreductase; NAD; Multigene family.	
FT	INIT MEM	0
FT	BINDING	151
FT	ACT SITE	178
FT	CONFLICT	224
SO	SEQUENCE	354 AA; 35322 MW; 6C6EC4ABAE54C377 CRC64;

```

RESULT      2
G3P2_AGABI
ID          G3P2_AGABI
AC          P32636.
              STANDARD;
              PRT;
              338 AA

```

SEQUENCE FROM N.A.

RA Wesels J.G.H.;

RT "Sequence analysis of the glyceraldehyde-3-phosphate dehydrogenase
RT genes from the basidiomycetes *Schizophyllum commune*, *Phanerochaete*
RT *chrysosporium* and *Agaricus bisporus*.",
RL *Curr. Genet.* 22:447-454(1992).
C 11- *Cantharellus baccatus*.

CC CRYSTALLINE ACTIVITY: D-GLYCEALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC + NAD(+) = 1,3-DIPHOSPHOGLYCERATE + NADH.
CC
CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC
CC -1- SUBUNIT: HOMOTETRAMER.
CC
CC -1- SUBCELLULAR LOCATION: CYTOSOL;MITOCHONDRIUM

-- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement.

 or send an email to license@lsb.slb.ch.
 (see <http://www.lsb-slb.ch/announce/>).

 EMBL: M81728; AAA32634.1; -.
 PIR: S26976; S26976.

DR HSSP; P00357; 1GPD.
DR InterPro; IPR000173; -.
DR Pfam; PF00044; gpdh; 1.
DR PRINTS; PR00078; G3PDHDCGNASE
DR PROSITE; PS00071; CADPU; 1

	GLYCOPOLYMERIZATION; NADH DEHYDROGENASE
ACT_SITE	177
SEQUENCE	338 AA; 36570 MW; 3381DBEF7B8A4469 CDSF.
	ACTIVATES THIOL DURING CATALYSIS

Query Match 62.5%; Score 65; DB 1; Length 338

Best Local Similarity 80.0%; Pred. No. 0.035;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVLVGVNGFGRIGRI 17
: ||:|||||
Db 1 MVKVGINGFGRIGRI 15

```

RESULT 3
G3P3_ECOLI STANDARD: PRT; 332 AA.
AC P33898;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE C (EC 1.2.1.12) (GAPDH-C).
GN GAPC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RA Hidalgo E., Limon A., Aguilar J.;
RL Submitted (XXX-1993) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 306-332 FROM N.A.
RC STRAIN-K12;
RA MEDLINE=92011371; PubMed=1917845;
RA Hidalgo E., Chen Y.-M., Lin E.C.C., Aguilar J.;
RT "Molecular cloning and DNA sequencing of the Escherichia coli K-12
RT ald gene encoding aldehyde dehydrogenase.";
RL J. Bacteriol. 173:6118-6123(1991).
RN [3]
RP SEQUENCE OF 1-13 FROM N.A.
RC STRAIN-K12;
RA MEDLINE=88232416; PubMed=2836696;
RA Nakamura H., Murakami H., Yamato I., Anraku Y.;
RT "Nucleotide sequence of the cydA gene encoding cytochrome b561 in
RT Escherichia coli K12.";
RL Mol. Gen. Genet. 212:1-5(1988).
RN [4]
RP CONCEPTUAL TRANSLATION.
RA Rudt K.E.;
RL Unpublished observations (JAN-1994).
CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC + NAD(+) -> 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; FRAMESHIFTS HAD TO
CC BE INTRODUCED TO PRODUCE THIS ORF.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL: L09067; AAA23856.1; ALT_FRAME.
CC EMBL: M64541; -; NOT_ANNOTATED_CDS.
CC EMBL: X07569; -; NOT_ANNOTATED_CDS.
CC HSSP: P17721; 1HOG.
CC Ecogen: B612103; gAPC.
CC InterPro: IPR000173; -.
CC Pfam: PF00044; gpdh; 1.
CC PRINTS: PR00078; G3PDHGNASE.
CC PROSITE: PS00071; GAPDH; 1.

```

KM Glycolysis; Oxidoreductase; NAD; Multigene family.
FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE
FT ACT_SITE 177 177 (BY SIMILARITY).
FT FT ACTIVATES THIOL GROUP DURING CATALYSIS
FT FT (BY SIMILARITY).
SQ SEQUENCE 332 AA; 35745 MW; C99B538697E001E CRC64;

Query Match 61.5%; Score 64; DB 1; Length 332;
Best Local Similarity 72.2%; Pred. No. 0.047;
Matches 13; Conservative 3; Mismatches 0; Indels 2; Gaps 1;

OY 6 VGVNGFGRIGR--IGRLV 21
||:|||||:|
Db 4 VGVNGFGRIGRLVGRLL 21

```

RESULT 4
G3P_COCHIE STANDARD: PRT; 337 AA.
AC P29497;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12) (GAPDH).
GN GPD1.
OS Cochliobolus heterostrophus (Drechslera maydis).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC Pleosporales; Pleosporaceae; Cochliobolus.
OX NCBI_Taxid=5016;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 48330 / C3;
RA MEDLINE=92306164; PubMed=1339326;
RA Van Wert S.L., Yoder O.C.;
RT "Structure of the Cochliobolus heterostrophus
RT glyceraldehyde-3-phosphate dehydrogenase gene.";
RL Curr. Genet. 22:28-35(1992).
CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC + NAD(+) -> 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL: X63516; CAA45084.1; -.
CC PIR: S22150; S22150.
CC PIR: S26946; S26946.
CC HSSP: P00357; 1GPD.
CC InterPro: IPR000173; -.
CC Pfam: PF00044; gpdh; 1.
CC PRINTS: PR00078; G3PDHGNASE.
CC PROSITE: PS00071; GAPDH; 1.
CC Glycolysis; Oxidoreductase; NAD.
CC ACT_SITE 178 178 GLYCERALDEHYDE 3-PHOSPHATE.
CC BINDING 151 151 ACTIVATES THIOL GROUP DURING CATALYSIS.
CC SEQUENCE 337 AA; 36544 MW; ABAF2743A362743E CRC64;

```

Query Match 61.5%; Score 64; DB 1; Length 337;
Best Local Similarity 80.0%; Pred. No. 0.048;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVLVGVNGFGRIGRI 17

- CATABOLIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE + NAD(+) = 1,3-DIPHOSPHATEGLYCERATE + NADH.
 - PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
 - SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 - SUBCELLULAR LOCATION: CYTOSOL.
 - SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: D88426; BAI13611.1; -.
CC HSSP: P00354; 3GPD.
CC InterPro: IPR000173; -.
CC Pfam: PF00044; gpdb: 1.
CC PRINTS: PR00078; G3PDHGNASE.
CC PROSITE: PS00071; GAPDH: 1.
CC KME Glycolysis: Oxidoreductase: NAD.
CC FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE.
CC FT ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS.
CC SO SEQUENCE 337 AA; 36724 MW; DFB4CB87CB4C1E0 CRC64;

Query Match 61.5%; Score 64; DB 1; Length 337;
Best Local Similarity 80.0%; Pred. No. 0.048;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVLVGNGFGRI 17
DB 1 VVVGNGFGRI 15

RESULT 8
G3P_NEUCR STANDARD; PRT; 338 AA.
ID G3P_NEUCR STANDARD; PRT; 338 AA.
AC P54118; Q92255;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12) (GAPDH)
DE (CLOCK-CONTROLLED PROTEIN 7).
DE GPD-1 OR CCG-7.
GN GPD-1 OR CCG-7.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Sahni M., Kinsey J.A.;
RT Identification and cloning of the Neurospora crassa glyceraldehyde-3-
RT phosphate dehydrogenase gene, gpd-1.";
RL Fungal Genet. NewsL. 44:47-49(1997).

RP SEQUENCE FROM N.A.
RA Dunlap J.C., Shinozuka M.L., Bell-Pedersen D., Ioros J.J.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC + NAD(+) = 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -1- SUBUNIT: HOMOTETRAMER (BT SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U56397; AAB00570.1; -.
CC DR EMBL: U67457; AAB95425.1; -.
CC DR HSSP: P00357; 1GPD.

DR InterPro: IPR000173; -.
DR Pfam: PF00044; gpdb: 1.
DR PRINTS: PR00078; G3PDHGNASE.
DR PROSITE: PS00071; GAPDH: 1.
CC KME Glycolysis: Oxidoreductase: NAD.
CC FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE
CC (BY SIMILARITY).
CC FT ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
CC (BY SIMILARITY).
CC FT CONFLICT 78 79 DA -> ER (IN REF. 2).
CC FT CONFLICT 212 212 A -> S (IN REF. 2).
CC SO SEQUENCE 338 AA; 36193 MW; 62CBF5896D3B2F57 CRC64;

Query Match 61.5%; Score 64; DB 1; Length 338;
Best Local Similarity 80.0%; Pred. No. 0.048;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVLVGNGFGRI 17
DB 2 VVVGNGFGRI 16

RESULT 9
G3PL_JACOR STANDARD; PRT; 363 AA.
ID G3PL_JACOR STANDARD; PRT; 363 AA.
AC P80534; Q64418;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, MUSCLE (EC 1.2.1.12)
DE (GAPDH).
OS Juncus orientalis.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Dipodidae; Dipodinae;
OC Juncus.
OX NCBI_TaxID=48868;
RN [1]
RP SEQUENCE OF 1-54.
RC TISSUE-Skeletal muscle;
RA Soukri A., Serrano A.;
RL Submitted (JAN-1996) to the SWISS-PROT data bank.

RP SEQUENCE OF 25-363 FROM N.A.
RC TISSUE-Skeletal muscle;
RX MEDLINE-97128782; PubMed-8973322;
RA Soukri A., Valverde F., Hafid N., Elkebdaj M.S., Serrano A.;
RT "Occurrence of a differential expression of the glyceraldehyde-3-
RT phosphate dehydrogenase gene in muscle and liver from eutheric and
RT induced hibernating jerboa (Juncus orientalis).";
RL Gene 181:139-145(1996).
CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC + NAD(+) = 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: X87226; GAA60678.1; -.
CC DR HSSP: P00354; 3GPD.
CC DR InterPro: IPR000173; -.
CC DR Pfam: PF00044; gpdb: 1.
CC DR PRINTS: PR00078; G3PDHGNASE.
CC PROSITE: PS00071; GAPDH: 1.

KW Glycolysis; Oxidoreductase; NAD.
 FT BINDING 180 180 GLYCERALDEHYDE 3-PHOSPHATE
 FT ACT_SITE 207 207 (BY SIMILARITY).
 FT SEQUENCE 363 AA; 39413 MW; 6CB549D9A1163C3F CRC64;
 (BY SIMILARITY).

Query Match 61.5%; Score 64; DB 1; Length 363;
 Best Local Similarity 80.0%; Pred. No. 0.051;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVLVGVNGFGRIGRI 17
 : |||||
 Db 1 MVKGVNGFGRIGRI 15

RESULT 10
 G3PC_MAIZE STANDARD; PRT; 337 AA.
 AC P08735;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, CYTOSOLIC 1 (EC 1.2.1.12).
 GN GAPC1 OR GAPC OR GPCL.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
 OC Andropogoneae; Zea.
 NCBI_TaxID=4577;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV. W22;
 RA MEDLINE=90040690; PubMed=2810356;
 RT Martinez P., Martin W.F., Cerff R.;
 RT "Structure, evolution and anaerobic regulation of a nuclear gene
 encoding cytosolic glyceraldehyde-3-phosphate dehydrogenase from
 maize";
 RL J. Mol. Biol. 208:551-565(1989).
 RN [2]
 RN SEQUENCE FROM N.A.
 RA MEDLINE=86230473; PubMed=3131533;
 RT Brinkmann H., Martinez P., Quigley F., Martin W.F., Cerff R.;
 RT "Endosymbiotic origin and codon bias of the nuclear gene for
 chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize";
 RL J. Mol. Evol. 26:320-328(1987).
 CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
 + NAD(+) = 1,3-DIPHOSPHATEGICERATE + NADH.
 CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- MISCELLANEOUS: PLANTS CONTAIN THREE FORMS OF GAPDH: A CYTOSOLIC
 FORM WHICH PARTICIPATES IN GLYCOLYSIS AND TWO CHLOROPLAST FORMS
 WHICH PARTICIPATES IN PHOTOSYNTHESIS. THESE THREE FORMS ARE
 ENCODED BY DISTINCT GENES.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, X07156; CA30151.1; -;
 DR EMBL, X15596; CA33620.1; -;
 DR PIR, S00354; DEZMGC.
 DR PIR, S06879; S06879.
 DR HSP, P00357; IGPD.
 DR MaltzDB, 13873; -;

DR Mendel; 383; ZEma;gapc.1.
 DR InterPro: IPR000173; -;
 DR Pfam: PF00044; gpdh.1.
 DR PRINTS: PR00078; G3PDHGNASE.
 DR PROSITE: PS00071; GAPDH.1.
 KW Glycolysis; Oxidoreductase; NAD; Multigene family.
 FT DOMAIN 1 151 BINDING TO NAD.
 FT DOMAIN 152 337 CATALYTIC.
 FT BINDING 154 154 GLYCERALDEHYDE 3-PHOSPHATE.
 FT ACT_SITE 181 181 ACTIVATES THIOL GROUP DURING CATALYSIS.
 FT CONFLICT 336 336 T->S (IN REF. 2).
 FT SEQUENCE 337 AA; 36522 MW; E18F580F09FDC07B CRC64;

Query Match 60.6%; Score 63; DB 1; Length 337;
 Best Local Similarity 62.5%; Pred. No. 0.066;
 Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 GLVLGVNGFGRIGRI 17
 : |||||
 Db 2 GKIKINGFGRIGRI 17

RESULT 11
 G3PC_MESCR STANDARD; PRT; 337 AA.
 AC P17878;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, CYTOSOLIC (EC 1.2.1.12).
 GN GAPC.
 OS Mesembryanthemum crystallinum (Common Ice Plant).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Eudicotyledons; core eudicots; Caryophyllidae;
 OC Caryophyllales; Alzooceae; Mesembryanthemum.
 NCBI_TaxID=3544;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA MEDLINE=90154012; PubMed=2303458;
 RT Ostrem J.A., Vernon D.M., Bohnert H.J.;
 RT "Increased expression of a gene coding for NAD:glyceraldehyde-3-
 phosphate dehydrogenase during the transition from C3 photosynthesis
 to crassulacean acid metabolism in Mesembryanthemum crystallinum";
 RL J. Biol. Chem. 265:3497-3502(1990).
 CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
 + NAD(+) = 1,3-DIPHOSPHATEGICERATE + NADH.
 CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- MISCELLANEOUS: PLANTS CONTAIN THREE FORMS OF GAPDH: A CYTOSOLIC
 FORM WHICH PARTICIPATES IN GLYCOLYSIS AND TWO CHLOROPLAST FORMS
 WHICH PARTICIPATES IN PHOTOSYNTHESIS. THESE THREE FORMS ARE
 ENCODED BY DISTINCT GENES.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, J05223; AA33033.1; -;
 DR EMBL, M29956; AA33031.1; -;
 DR PIR, A35080; A35080.
 DR HSP, P00357; IGPD.
 DR Mendel; 74; MESCR;gapc.1.
 DR InterPro: IPR000173; -;
 DR Pfam, PF00044; gpdh.1.
 DR PRINTS, PR00078; G3PDHGNASE.

DR PROSITE: PS00071: GAPDH: 1.
 KM Glycolysis; Oxidoreductase; NAD; Multigene family.
 FT DOMAIN 1 151 BINDING TO NAD.
 FT BINDING 152 337 CATALYTIC.
 FT BINDING 154 154 GLYCERALDEHYDE 3-PHOSPHATE.
 FT ACT_SITE 181 181 ACTIVATES THIOL GROUP DURING CATALYSIS.
 SO SEQUENCE 337 AA: 36587 MW: B07B8A25509E9EA CRC64;

Query Match 60.6%; Score 63; DB 1; Length 337;
 Best Local Similarity 68.4%; Pred. No. 0.066;
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 VLVGVNGFGRIGRIGRLV 22
 DB 4 VLVGVNGFGRIGRIGRLV 22

RESULT 12
 G3PC_ORYSA

ID G3PC_ORYSA STANDARD; PRT: 337 AA.

AC 042977;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, CYTOSOLIC (EC 1.2.1.12).
 GN GAPC OR GPC.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae;
 OC Oryza.
 OX NCBI_TaxID=4530;

RP SEQUENCE FROM N.A.
 RC STRAIN=CV. INDICA GUANG-LU-AI NO.4; TISSUE=Shoot;

RA Xiao C., Jiang Y., Cao K.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
 + NAD(+) -> 1,3-DIPHOSPHATEGLYCERATE + NADH.

CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.

CC -1- SUBUNIT: HOMOTETRAMER.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -1- MISCELLANEOUS: PLANTS CONTAIN THREE FORMS OF GAPDH: A CYTOSOLIC
 FORM WHICH PARTICIPATES IN GLYCOLYSIS AND TWO CHLOROPLAST FORMS
 WHICH PARTICIPATES IN PHOTOSYNTHESIS. THESE THREE FORMS ARE

CC ENCODED BY DISTINCT GENES.

CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 DEHYDROGENASE FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

DR EMBL: U31676; AA82047.1; -
 DR HSP: P00357; IGPD.
 DR Mendel: 9724; ORYSA:GAPC.1.
 DR InterPro: IPR000173; -
 DR Pfam: PF00044; gpdh; 1.
 DR PRINTS: PR00078; G3PDHGNASE.
 DR PROSITE: PS00071; GAPDH; 1.
 KM GLYCOLYSIS; Oxidoreductase; NAD; Multigene family.
 FT BINDING 154 154 GLYCERALDEHYDE 3-PHOSPHATE.
 FT ACT_SITE 181 181 ACTIVATES THIOL GROUP DURING CATALYSIS.
 SO SEQUENCE 337 AA: 36492 MW: ABB5F8B64E3B35F CRC64;

Query Match 60.6%; Score 63; DB 1; Length 337;
 Best Local Similarity 62.5%; Pred. No. 0.066;
 Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 GLVGVNGFGRIGRIGRL 17
 DB 2 GKIKINGFGRIGRL 17

RESULT 13
 G3PD_MAIZE

ID G3PD_MAIZE STANDARD; PRT: 337 AA.

AC 009054;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, CYTOSOLIC 2 (EC 1.2.1.12).
 GN GAPC2 OR GPC2.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
 OC Andropogoneae; Zea.
 OX NCBI_TaxID=4577;

RP SEQUENCE FROM N.A.
 RC STRAIN=CV. RP704;
 RA Liand M.-F.;
 RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=CV. B73;
 RA Manjunath S., Sachs M.M.;
 RT Molecular characterization and promoter analysis of the maize
 cytosolic glyceraldehyde 3-phosphate dehydrogenase gene family and its
 expression during anoxia.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE OF 91-337 FROM N.A.
 RC STRAIN=CV. BERKELEY FAST; TISSUE=coleoptile;
 RX MEDLINE=92393412; PubMed=2535522;
 RA Russell D.A., Sachs M.M.;
 RL Plant Cell 1:793-803(1989).
 RT Differential expression and sequence analysis of the maize
 glyceraldehyde-3-phosphate dehydrogenase gene family.;

CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
 + NAD(+) -> 1,3-DIPHOSPHATEGLYCERATE + NADH.

CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.

CC -1- SUBUNIT: HOMOTETRAMER.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -1- TISSUE SPECIFICITY: DEVELOPING SEED, SEEDLING ROOTS AND SHOOTS,
 AND EMBRYO.

CC -1- DEVELOPMENTAL STAGE: EXPRESSION DECREASES DURING SEED DEVELOPMENT
 AND EMBRYO MATURATION.

CC -1- MISCELLANEOUS: PLANTS CONTAIN THREE FORMS OF GAPDH: A CYTOSOLIC
 FORM WHICH PARTICIPATES IN GLYCOLYSIS AND TWO CHLOROPLAST FORMS
 WHICH PARTICIPATES IN PHOTOSYNTHESIS. THESE THREE FORMS ARE

CC ENCODED BY DISTINCT GENES.

CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 DEHYDROGENASE FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

DR EMBL: U45858; AA87880.1; -
 DR EMBL: X73151; CAA51676.1; -
 DR EMBL: U45855; AA87578.1; -
 DR EMBL: L13432; AAA33466.1; -
 DR Mendel: 575; ZEAMA:GAPC.2.
 DR InterPro: IPR000173; -
 DR Pfam: PF00044; gpdh; 1.
 DR PRINTS: PR00078; G3PDHGNASE.

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:29:46 ; Search time 150.43 Seconds
(Without alignments)
17.141 Million cell updates/sec

Title: PCT-US01-05825A-31
Perfect score: 104
Sequence: 1 XGLVLGVNGVGFGRIGRILVI 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_15: *
2: sp_archaea: *
3: sp_bacteria: *
4: sp_fungi: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_unclassified: *
13: sp_vertebrate: *
14: sp_virus: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	61.5	165	11 003850	003850 rattus norv
2	64	61.5	333	6 09N2D6	09N2D6 canis famli
3	64	61.5	333	6 09N2D5	09N2D5 felis silve
4	64	61.5	333	11 090WU4	090WU4 rattus norv
5	64	61.5	335	3 09UW96	09UW96 pleurotus s
6	64	61.5	337	3 09P8C0	09P8C0 phaseosphaer
7	64	61.5	339	3 09Y8E9	09Y8E9 cryptococcu
8	63	60.6	333	13 09PTW5	09PTW5 paratrichthy
9	63	60.6	337	3 09UR38	09UR38 leuticula e
10	63	60.6	338	3 09UV22	09UV22 ptafilia rho
11	62	59.6	39	2 09RSJ2	09RSJ2 streptococc
12	62	59.6	75	6 09T6S3	09T6S3 equus cabal
13	62	59.6	76	6 09T6S2	09T6S2 felis silve
14	62	59.6	106	6 09M205	09M205 ovis aries
15	61.5	59.1	87	10 041949	041949 arabidopsis
16	61.5	59.1	338	10 09M8W8	09M8W8 arabidopsis
17	61.5	59.1	340	10 004231	004231 selaginella
18	61	58.7	105	3 09Y7H1	09Y7H1 pitomyces s
19	61	58.7	334	3 09UVCO	09UVCO pichia cife

20	61	58.7	337	3 09Y796	09Y796 cryptococcu
21	61	58.7	339	10 09M7R9	09M7R9 achlya bisce
22	61	58.7	368	10 06S843	06S843 marsilia qu
23	61	58.7	422	10 09SAU6	09SAU6 arabidopsis
24	60	57.7	17	6 09T0Z5	09T0Z5 macaca fasc
25	60	57.7	49	5 P90514	P90514 acanthameob
26	60	57.7	83	11 P97617	P97617 rattus norv
27	60	57.7	333	2 09KT36	09KT36 vlbrio chol
28	60	57.7	337	10 09XG67	09XG67 nicotiana t
29	60	57.7	338	2 032755	032755 lactobacill
30	60	57.7	361	5 096424	096424 herpetomona
31	59	56.7	32	5 09NUT5	09NUT5 leishmania
32	59	56.7	320	2 0915X6	0915X6 streptococc
33	59	56.7	330	2 09NUP1	09NUP1 delnococtus
34	59	56.7	333	2 09WML7	09WML7 synecococc
35	59	56.7	337	10 043359	043359 zea mays (m
36	59	56.7	338	5 001373	001373 schistosoma
37	59	56.7	338	5 027776	027776 schistosoma
38	59	56.7	339	5 09N655	09N655 mastigameob
39	59	56.7	353	10 043311	043311 euglena gra
40	59	56.7	361	5 096423	096423 crithidia f
41	59	56.7	363	5 026753	026753 trypanoplas
42	59	56.7	363	5 026754	026754 trypanoplas
43	59	56.7	614	10 09LE93	09LE93 phaeodactyl
44	59	56.7	615	10 09M7R5	09M7R5 odontella s
45	58.5	56.2	388	10 004167	004167 pyrenomonas

ALIGNMENTS

RESULT 1
003850 PRELIMINARY: PRT: 165 AA.
ID 003850; 063208;
AC 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (NADP+) (EC 1.2.1.13)
DE (TRIOSEPHOSPHATE DEHYDROGENASE (NADP+)) (RAT GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE) (FRAGMENT).
GN G3PD.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE DAWLEY; TISSUE=TESTIS;
RA Kravetz S.;
RL Submitted (Oct-1990) to the EMBL/GenBank/DBJ databases.
RN [12]
RP SEQUENCE OF 1-52 FROM N.A.
RX MEDLINE=87053179; PubMed=3780374;
RA Kravetz S.A.; Connor W.; Cannon P.D.; Dixon G.H.;
RT Libraries: evidence for a rat glyceraldehyde-3-phosphate dehydrogenase-like mRNA and a ferritin mRNA within testis.";
RL DNA 5:427-435(1986).
CC -I- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + NADP(+) + NAD(+) -> 3-PHOSPHO-D-GLYCEROL PHOSPHATE + NADPH.
DR EMBL: X54798; CA38569.1; -;
DR EMBL: M14166; AAA41178.1; -;
DR HSSP: P17721; IHG.
DR INTERPRO: IPR000173; -;
DR INTERPRO: IPR000531; -;
DR PFAM: PF00044; gpdh. 1.
DR PROSITE: PS00073; GAPDH. 1.
DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW NAD; NADP.
FT NON_TER
SQ SEQUENCE 165 AA; 17737 MW; EAE35E4B8E705A02 CRC64;

Query Match 61.5%; Score 64; DB 11; Length 165;
 Best Local Similarity 72.2%; Pred. No. 0.086;
 Matches 13; Conservative 3; Mismatches 0; Indels 2; Gaps 1;

OY 6 VGVNGFGRIGR--IGRLV 21
 : ||:||||||| :|||:
 Db 4 VGINFGFGRIGRLVGRLL 21

RESULT 2

ID 09N2D6 PRELIMINARY; PRT; 333 AA.

AC 09N2D6; 01-OCT-2000 (TREMBLREL. 15, Created)
 DT 01-OCT-2000 (TREMBLREL. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLREL. 15, Last annotation update)
 DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE.
 GN GAPDH.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ADRENAL GLAND;
 RA Kitamura H., Adachi K., Kido N., Hagiya T., Minase K., Yasui H.,
 Yano E., Ohta Y., Tabu K., Mae J., Kanehira K., Ohashi A.;
 RT "Canine glyceraldehyde-3-phosphate dehydrogenase (GAPDH), complete
 CDS."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB038240; BAA90817.1; -
 KW SEQUENCE 333 AA; 35861 MW; E08B4C39FA2A7EE CRC64;

Query Match 61.5%; Score 64; DB 6; Length 333;
 Best Local Similarity 80.0%; Pred. No. 0.19;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVLVGVNGFGRIGRI 17
 : ||||||||| :|||:
 Db 1 MVKGVNGFGRIGRL 15

RESULT 3

ID 09N2D5 PRELIMINARY; PRT; 333 AA.

AC 09N2D5; 01-OCT-2000 (TREMBLREL. 15, Created)
 DT 01-OCT-2000 (TREMBLREL. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLREL. 15, Last annotation update)
 DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE.
 GN GAPDH.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-WHITE ADIPOSE TISSUE;
 RA Kitamura H., Adachi K., Ohta Y., Kido N., Hagiya T., Yasui H.,
 Yano E., Minase K., Mae J., Tabu K., Kanehira K., Ohashi A.;
 RT "Feline glyceraldehyde-3-phosphate dehydrogenase (GAPDH), complete
 CDS."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB038241; BAA90818.1; -
 KW SEQUENCE 333 AA; 35813 MW; 20501C401BBE7906 CRC64;

Query Match 61.5%; Score 64; DB 6; Length 333;
 Best Local Similarity 80.0%; Pred. No. 0.19;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVLVGVNGFGRIGRI 17
 : ||||||||| :|||:
 Db 1 MVKGVNGFGRIGRL 15

RESULT 4

ID 09QWU4 PRELIMINARY; PRT; 333 AA.

AC 09QWU4; 01-MAY-2000 (TREMBLREL. 13, Created)
 DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLREL. 15, Last annotation update)
 DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12).
 GN GAPDH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-BRAIN;
 RA Zheng J., Ramirez V.D.;
 RT "Cloning of a rat brain glyceraldehyde-3-phosphate dehydrogenase cDNA
 by a rapid PCR-based method."
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF106860; AAD08929.1; -
 DR HSSP: P00354; 3GPD.
 DR INTERPRO: IPR000173; -
 DR PFAM: PF00044; gpdh; 1.
 DR PRINTS: PR00078; G3PDHROGNASE.
 DR PROSITE: PS00071; GAPDH; 1.
 KW Oxidoreductase.
 KW SEQUENCE 333 AA; 35844 MW; 3185EADAA544180 CRC64;

Query Match 61.5%; Score 64; DB 11; Length 333;
 Best Local Similarity 80.0%; Pred. No. 0.19;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVLVGVNGFGRIGRI 17
 : ||||||||| :|||:
 Db 1 MVKGVNGFGRIGRL 15

RESULT 5

ID 09UW96 PRELIMINARY; PRT; 335 AA.

AC 09UW96; 01-MAY-2000 (TREMBLREL. 13, Created)
 DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLREL. 15, Last annotation update)
 DE GLYCEROL-3-PHOSPHATE DEHYDROGENASE (FRAGMENT).
 OS Pleurotus sajor-caju (Oyster mushroom).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;
 OC Pleurotaceae; Pleurotus.
 OX NCBI_TaxID=50053;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ASI2070; TISSUE-MYCELIUM;
 RA Jeong M.-J., Park S.-C., Byun M.-O., Ryu J.-C.;
 RT "Isolation and characterization of the gene encoding glycerol-3-
 phosphate dehydrogenase from Pleurotus sajor-caju."
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF087676; AAD52091.1; -
 DR HSSP: P56649; IDSS.
 DR INTERPRO: IPR000173; -
 DR PFAM: PF00044; gpdh; 1.
 DR PRINTS: PR00078; G3PDHROGNASE.
 DR NON_TER 335
 KW SEQUENCE 335 AA; 36065 MW; 384C26ADF341A841 CRC64;

Query Match 61.5%; Score 64; DB 3; Length 335;

```

Best Local Similarity 80.0%; Pred. No. 0.19;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVLVGNFGFRIGRI 17
Db 1 MVKVGINGFRIGRI 15

RESULT 6
Q9P8C0 PRELIMINARY; PRT; 337 AA.
AC Q9P8C0:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12).
GN GPD1.
OS Phaeosphaeria nodorum (Septoria nodorum).
OC Eukaryota; Fungi; Ascomycota; Ascomycota Incertae sedis;
OC Phaeosphaeriaceae; Phaeosphaeria.
OX NCBI_TaxID=13684;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=BSM300;
RA Cooley R.N.;
RT "The glyceraldehyde 3-phosphate dehydrogenase gene of Stagonospora
RT nodorum."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ271155; CAB72263.1; -.
KM Oxidoreductase.
SQ SEQUENCE 337 AA; 36491 MW; 3206BADA5728DF5 CRC64;

Query Match 61.5%; Score 64; DB 3; Length 337;
Best Local Similarity 80.0%; Pred. No. 0.2;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVLVGNFGFRIGRI 17
Db 2 VVKVGINGFRIGRI 16

RESULT 7
Q9Y8E9 PRELIMINARY; PRT; 339 AA.
AC Q9Y8E9:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12).
GN GPD.
OS Cryptococcus neoformans (Filobasidiella neoformans).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Tremellales;
OC Tremellaceae; Filobasidiella.
OX NCBI_TaxID=5207;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=B-3501;
RX MEDLINE=99282520; PubMed=10352226;
RA Varma A.K., Kwon-Chung K.J.;
RT "Characterization of the glyceraldehyde-3-phosphate gene and the use
RT of its promoter for heterologous expression in Cryptococcus
RT neoformans, a human pathogen."
RL Gene 232:155-163(1999).
CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC + NAD(+) = 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC EMBL: AF106950; AAD29256.1; -.
DR HSSP: P00354; 3GPD.
DR INTERPRO: IPR000173; -.
DR PFAM: PF00044; gpdh; 1.
DR PRINTS: PR00078; G3PDHRCGNASE.
DR PROSITE: PS00071; GAPDH; 1.

```

```

KM Oxidoreductase; NAD.
SQ SEQUENCE 339 AA; 36351 MW; 1F1459F3A7D1B8FF CRC64;

Query Match 61.5%; Score 64; DB 3; Length 339;
Best Local Similarity 80.0%; Pred. No. 0.2;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVLVGNFGFRIGRI 17
Db 2 VVKVGINGFRIGRI 16

RESULT 8
Q9PTW5 PRELIMINARY; PRT; 333 AA.
AC Q9PTW5:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (GAPDH).
OS Parailichthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorph; Pleuronectiformes;
OC Pleuronectoidel; Boichidae; Parailichthys.
OX NCBI_TaxID=8255;
RN 11
RP SEQUENCE FROM N.A.
RA Aoki T., Naka H., Katagiri T., Hiroo I.;
RT "Cloning and characterization of glyceraldehyde-3-phosphate
RT dehydrogenase (GAPDH) cDNA of Japanese flounder, Parailichthys
RT olivaceus."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB029337; BAA8638.1; -.
DR HSSP: P00354; 3GPD.
DR INTERPRO: IPR000173; -.
DR PFAM: PF00044; gpdh; 1.
DR PRINTS: PR00078; G3PDHRCGNASE.
DR PROSITE: PS00071; GAPDH; 1.
SQ SEQUENCE 333 AA; 36041 MW; CED70B2E082B8757 CRC64;

Query Match 60.6%; Score 63; DB 13; Length 333;
Best Local Similarity 73.3%; Pred. No. 0.27;
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVLVGNFGFRIGRI 17
Db 1 MVKVGINGFRIGRI 15

RESULT 9
Q9UR38 PRELIMINARY; PRT; 337 AA.
AC Q9UR38:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12).
GN GLEGPD.
OS Lentinula edodes (Shiitake mushroom) (Lentinus edodes).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;
OC Tricholomataceae; Lentinula.
OX NCBI_TaxID=5353;
RN 11
RP SEQUENCE FROM N.A.
RA Hirano T., Sato T., Okawa K., Kanda K., Yaegashi K., Enel H.;
RT "Isolation and characterization of the glyceraldehyde-3-phosphate
RT dehydrogenase gene of Lentinus edodes."
RL Biosci. Biotechnol. Biochem. 63:1223-1227(1999).
DR EMBL: AB013136; BAA83550.1; -.
DR EMBL: AB012862; BAA83549.1; -.

```


DR INTERPRO: IPR000173; -
 DR PRAM: PF00044; gpdt: 1.
 FT NON_TER 1
 FT NON_TER 76
 SO SEQUENCE 76 AA; 8334 MW; D3F3E09146FF7936 CRC64;

Query Match 59.6%; Score 62; DB 6; Length 76;
 Best Local Similarity 85.7%; Pred. No. 0.068;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 VLVGVNGFGRIGRI 17
 :|:|||||:
 DB 3 VKGVNGFGRIGRI 16

RESULT 14
 09M205 PRELIMINARY; PRT: 106 AA.

AC 09M205;
 DT 01-OCT-2000 (TREMBLER. 15, Created)
 DT 01-OCT-2000 (TREMBLER. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLER. 15, Last annotation update)
 DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE GAPDH (FRAGMENT).
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Herrmann L.M., Davis W.C., Li H., Wardrup J., Sy M.-S., Gambetti P.,
 RA Knowles D.P.;
 RT "Detection of PrPC in peripheral blood mononuclear cells of normal
 RT sheep."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF272837; AAF91410.1; -
 FT NON_TER 1
 FT NON_TER 106
 SO SEQUENCE 106 AA; 11612 MW; B22E924D623D550A CRC64;

Query Match 59.6%; Score 62; DB 6; Length 106;
 Best Local Similarity 85.7%; Pred. No. 0.1;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 VLVGVNGFGRIGRI 17
 :|:|||||:
 DB 1 VKGVNGFGRIGRI 14

RESULT 15
 041949 PRELIMINARY; PRT: 87 AA.

AC 041949;
 DT 01-NOV-1996 (TREMBLER. 01, Created)
 DT 01-NOV-1996 (TREMBLER. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBLER. 15, Last annotation update)
 DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (FRAGMENT).
 GN GAPC.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-WHOLE SEEDLINGS OF A. THALIANA ECOTYPE COLUMBIA;
 RA Desprez T., Amsalem J., Chiapello H., Rouze P., Caboche M., Hofte H.;
 RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
 DR HSSP: P00354; 3GPD.
 DR MENDEL: 7185; Arabid.GAPC; 7185.

DR INTERPRO: IPR000173; -
 DR PRAM: PF00044; gpdt: 1.
 KW Multigene family; Oxidoreductase.
 FT NON_TER 87
 SO SEQUENCE 87 AA; 10034 MW; ADC06618BA8F5251 CRC64;

Query Match 59.1%; Score 61.5; DB 10; Length 87;
 Best Local Similarity 55.0%; Pred. No. 0.094;
 Matches 11; Conservative 6; Mismatches 2; Indels 1; Gaps 1;

OY 4 VLVGVNGFGRIGRI-IGRLVI 22
 :|:|||||:|:|:
 DB 6 IRIGNGFGRIGRIIVARYVL 25

Search completed: June 13, 2001, 14:29:47
 Job time: 548 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:27:09 ; Search time 78.71 Seconds
(without alignments)
5.370 Million cell updates/sec

Title: PCT-US01-05825A-31
Sequence: 1 XGLVGVNGRGRIGRIGRLVI 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents-AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTC0S.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	59.6	18	1	US-08-553-110-1
2	62	59.6	18	3	US-08-946-026-50
3	62	59.6	333	4	US-08-961-083-54
4	62	59.6	336	1	US-07-928-462-2
5	62	59.6	336	4	US-08-273-247-2
6	59	58.7	268	2	US-08-997-080-186
7	59	58.7	268	4	US-08-997-362-186
8	59	56.7	268	2	US-09-095-855-186
9	59	56.7	334	6	5290690-10
10	59	56.7	334	6	5290690-11
11	59	56.7	335	2	US-08-903-800A-6
12	59	56.7	335	6	5290690-9
13	59	56.7	340	4	US-09-095-855-205
14	55	52.9	12	2	US-08-464-073-33
15	52	50.0	27	6	5245013-20
16	50	48.1	22	1	US-08-678-444-1
17	50	48.1	416	4	US-09-007-476-2
18	48	46.2	277	3	US-09-007-484-4
19	48	46.2	425	2	US-08-986-963-2
20	48	46.2	430	3	US-09-007-484-2
21	43	41.3	22	1	US-08-678-444-2
22	41	39.4	345	3	US-09-222-817-2
23	41	39.4	530	3	US-09-222-817-12
24	41	39.4	530	3	US-09-222-817-14
25	39	37.5	327	1	US-08-240-049B-14
26	39	37.5	327	1	US-08-259-148A-16
27	39	37.5	327	1	US-08-484-054-16

28	39	37.5	327	2	US-07-876-941A-16	Sequence 16, Appl
29	39	37.5	327	3	PCT-US95-13703-18	Sequence 18, Appl
30	39	37.5	435	1	US-08-259-148A-18	Sequence 18, Appl
31	39	37.5	435	1	US-08-484-054-18	Sequence 18, Appl
32	39	37.5	435	2	US-07-876-941A-18	Sequence 18, Appl
33	39	37.5	525	5	PCT-US95-13703-28	Sequence 28, Appl
34	39	37.5	540	5	PCT-US95-13703-26	Sequence 26, Appl
35	39	37.5	549	5	PCT-US95-13703-16	Sequence 16, Appl
36	39	37.5	659	1	US-08-259-148A-16	Sequence 16, Appl
37	39	37.5	659	1	US-08-484-054-20	Sequence 20, Appl
38	39	37.5	659	2	US-07-876-941A-20	Sequence 20, Appl
39	39	37.5	659	2	PCT-US95-13703-14	Sequence 14, Appl
40	39	37.5	660	1	US-07-803-633A-13	Sequence 13, Appl
41	38.5	37.0	865	5	US-08-602-010A-16	Sequence 16, Appl
42	38	36.5	152	1	US-08-680-726A-16	Sequence 16, Appl
43	38	36.5	152	4	US-09-092-409-16	Sequence 16, Appl
44	38	36.5	305	1	US-08-680-726A-80	Sequence 80, Appl
45	38	36.5				

ALIGNMENTS

RESULT 1
US-08-553-110-1
Sequence 1, Application US/08553110
Patent No. 5723301
GENERAL INFORMATION:
APPLICANT: Burke, James R.
APPLICANT: Vance, Jeffrey M.
APPLICANT: Engchild, Jan
APPLICANT: Stittmattner, Warren J.
TITLE OF INVENTION: Therapeutics for Diseases Associated
TITLE OF INVENTION: with CAG trinucleotide Repeat Expansion
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Kenneth D. Sibley
STREET: P.O. Drawer 34009
CITY: Charlotte
STATE: No. 5723301th Carolina
COUNTRY: United States of America
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/553,110
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5405-117
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-553-110-1
Query Match 59.6%; Score 62; DB 1; Length 18;
Best Local Similarity 85.7%; Pred. No. 0.0012;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 4 VLVGVNGRGRIGRI 17

Db 1 VKVGVNGFGRIGRL 14

```
RESULT 2
US-08-946-026-50
: Sequence 50, Application US/08946026
: Patent No. 6034218
: GENERAL INFORMATION:
: APPLICANT: Reed, Steven G.
: APPLICANT: Dillon, Davin C.
: APPLICANT: Twardzik, Daniel R.
: APPLICANT: Mitcham, Jennifer L.
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
: TITLE OF INVENTION: AND IMMUNODIAGNOSIS OF PROSTATE CANCER
: NUMBER OF SEQUENCES: 59
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SEED and BERRY LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98104-7092
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/946,026
: FILING DATE: 07-OCT-1997
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Makl, David J.
: REGISTRATION NUMBER: 31,382
: REFERENCE/DOCKET NUMBER: 210121,424C1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 50:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 18 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-946-026-50

Query Match 59.6%; Score 62; DB 3; Length 18;
Best Local Similarity 85.7%; Pred. No. 0.0012;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VLKGVNGFGRIGRI 17
Db 3 VKVGVNGFGRIGRL 16

RESULT 3
US-08-961-083-54
: Sequence 54, Application US/08961083
: Patent No. 6159469
: GENERAL INFORMATION:
: APPLICANT: Choi et. al.
: TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
: NUMBER OF SEQUENCES: 452
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
```

```
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,083
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 54:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 333 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-961-083-54
```

```
Query Match 59.6%; Score 62; DB 4; Length 333;
Best Local Similarity 73.3%; Pred. No. 0.032;
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVLKGVNGFGRIGRI 17
Db 1 VKVGVNGFGRIGRL 15
```

```
RESULT 4
US-07-928-462-2
: Sequence 2, Application US/07928462
: Patent No. 5328996
: GENERAL INFORMATION:
: APPLICANT: Boyle, Michael D.P.
: APPLICANT: Lotenberg, Richard
: APPLICANT: Broder, Christopher C.
: APPLICANT: von Merling, Gregory O.
: TITLE OF INVENTION: Bacterial Plasmid Receptors as
: TITLE OF INVENTION: Fibriolytic Agents
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: David R. Saliwanchik
: STREET: 2421 N.W. 41st Street, Suite A-1
: CITY: Gainesville
: STATE: FL
: COUNTRY: USA
: ZIP: 32606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/928,462
: FILING DATE: 19920810
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/524,411
: FILING DATE: 16-May-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/330,849
: FILING DATE: 29-MAR-1989
: ATTORNEY/AGENT INFORMATION:
: NAME: Saliwanchik, David R.
```

REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF/SAS-13.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 336 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-928-462-2

Query Match 59.6%; Score 62; DB 1; Length 336;
Best Local Similarity 73.3%; Pred. No. 0.032;
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVLVGNFGFRIGRI 17
: |||:|||||:
Db 2 VVKVGNFGFRIGRI 16

RESULT 5
US-08-273-247-2
Sequence 2, Application US/08273247
Patent No. 6136323
GENERAL INFORMATION:
APPLICANT: Boyle, Michael D.P.
APPLICANT: Lottenberg, Richard
APPLICANT: Broder, Christopher C.
APPLICANT: von Merling, Gregory O.
TITLE OF INVENTION: Bacterial Plasmid Receptors as
NUMBER OF SEQUENCES: 2
TITLE OF INVENTION: Fibrinolytic Agents
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Sallwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/273,247
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/524,411
FILING DATE: 16-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/330,849
FILING DATE: 29-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sallwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF/SAS-13.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 336 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-273-247-2

Query Match 59.6%; Score 62; DB 4; Length 336;
Best Local Similarity 73.3%; Pred. No. 0.032;
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 LVLVGNFGFRIGRI 17
: |||:|||||:
Db 2 VVKVGNFGFRIGRI 16

RESULT 6
US-08-997-080-186
Sequence 186, Application US/08997080
Patent No. 5968524
GENERAL INFORMATION:
APPLICANT: WATSON, JAMES D.
APPLICANT: TAN, PAUL L.J.
TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,080
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 186:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-080-186

Query Match 56.7%; Score 59; DB 2; Length 268;
Best Local Similarity 84.6%; Pred. No. 0.072;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 LVLVGNFGFRIGRI 16
: |||:|||||:
Db 3 IRVGVNFGFRIGRI 15

RESULT 7
US-08-997-362-186
Sequence 186, Application US/08997362
Patent No. 5985287
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Hiyama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Scott, Linda

APPLICANT: Prestidge, Ross
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,362
CLASSIFICATION:
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
FILING DATE: June 12, 1997
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
FILING DATE: August 29, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 186:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-362-186

Query Match 56.7%; Score 59; DB 2; Length 268;
Best Local Similarity 84.6%; Pred. No. 0.072;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 VLVGVNGFGRIGR 16
: |||||
Db 3 IRVGVNGFGRIGR 15

RESULT 8
US-09-095-855-186
Sequence 186, Application US/09095855
GENERAL INFORMATION:
PATENT NO. 6160093
APPLICANT: Tan, Paul
APPLICANT: Vliesser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 186:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-095-855-186

Query Match 56.7%; Score 59; DB 4; Length 268;
Best Local Similarity 84.6%; Pred. No. 0.072;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 VLVGVNGFGRIGR 16
: |||||
Db 3 IRVGVNGFGRIGR 15

RESULT 9
5290690-10
PATENT NO. 5290690
APPLICANT: MRABET, NADIR;LASTERS, IGNACE;STANSENS, PATRICK
;MATTHYSSENS, GASTON;WODAK, SHOSHANA;QUAX, WILHELMUS J.
TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
;STABILITY OF PROTEINS
NUMBER OF SEQUENCES: 22
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/398,706
FILING DATE: 25-AUG-1989
SEQ ID NO:10
LENGTH: 334
5290690-10

Query Match 56.7%; Score 59; DB 6; Length 334;
Best Local Similarity 84.6%; Pred. No. 0.092;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 VLVGVNGFGRIGR 16
: |||||
Db 2 VKVGVNGFGRIGR 14

RESULT 10
5290690-11
PATENT NO. 5290690
APPLICANT: MRABET, NADIR;LASTERS, IGNACE;STANSENS, PATRICK
;MATTHYSSENS, GASTON;WODAK, SHOSHANA;QUAX, WILHELMUS J.
TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
;STABILITY OF PROTEINS
NUMBER OF SEQUENCES: 22
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/398,706
FILING DATE: 25-AUG-1989
SEQ ID NO:11
LENGTH: 334
5290690-11

Query Match 56.7%; Score 59; DB 6; Length 334;
Best Local Similarity 84.6%; Pred. No. 0.092;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VLVGVNGFGRRGR 16
|11:|||||
DB 2 VKVINGFGRRGR 14

RESULT 11
US-08-903-800A-6
Sequence 6, Application US/08903800A
GENERAL INFORMATION:
APPLICANT: RHEE, Sang-Ki
APPLICANT: CHOI, Eun-Sung
APPLICANT: KIM, Chul-Ho
APPLICANT: SOHN, Jung-Hoon
APPLICANT: KANG, Hyun-Ah
APPLICANT: KIM, Hwa-Young
TITLE OF INVENTION: AUTONOMOUSLY REPLICATING SEQUENCES,
TITLE OF INVENTION: GAPDH GENE AND PROMOTER DERIVED FROM
TITLE OF INVENTION: HANSENULA POLYMORPHA, EXPRESSION
TITLE OF INVENTION: VECTORS CONTAINING SAME AND METHOD FOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: RHEE, Sang-Ki
STREET: Keukdong Villa Ka-101, Kwangjang-dong,
STREET: Kwangjin-gu
CITY: Seoul
STATE: Seoul
COUNTRY: Republic of Korea
ZIP: 143-210
ADDRESSEE: CHOI, Eun-Sung
STREET: Dasol Apt. 102-507, #395-3, Kung-dong,
STREET: Yuseong-gu
CITY: Daejeon
STATE: Daejeon
COUNTRY: Republic of Korea
ZIP: 305-335
ADDRESSEE: KIM, Chul-Ho
STREET: Shindonga Apt. 2-1101, Yongjeon-dong, Dong-gu
CITY: Daejeon
STATE: Daejeon
COUNTRY: Republic of Korea
ZIP: 300-200
ADDRESSEE: SOHN, Jung-Hoon
STREET: Nuri Apt. 103-506, Wolpyung-dong, Seo-gu
CITY: Daejeon
STATE: Daejeon
COUNTRY: Republic of Korea
ZIP: 302-280
ADDRESSEE: KANG, Hyun-Ah
STREET: Kyungseong Keunmaul Apt. 125-1501, Kalma-dong,
STREET: Seo-gu
CITY: Daejeon
STATE: Daejeon
COUNTRY: Republic of Korea
ZIP: 302-171
ADDRESSEE: KIM, Hwa-Young
STREET: #3-462, Bukahyun-3-dong, Seodaemun-gu
CITY: Seoul
STATE: Seoul
COUNTRY: Republic of Korea
ZIP: 120-190

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5inch, 1.44MB storage
COMPUTER: IBM PC/AT
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/903,800A
FILING DATE: 31-JUL-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: KR 97-3173
FILING DATE: 31-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE:
TELEFAX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 335 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
OTHER INFORMATION: glyceraldehyde-3-phosphate
OTHER INFORMATION: dehydrogenase of Hansenula polymorpha
US-08-903-800A-6

Query Match 56.7%; Score 59; DB 2; Length 335;
Best Local Similarity 83.3%; Pred. No. 0.093;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 VGVNGFGRRGR 17
|1:|||||
DB 5 VGVNGFGRRGR 16

RESULT 12
5290690-9
Patent No. 5290690
APPLICANT: MRABET, NADIR;LASTERS, IGNACE;STANSSENS, PATRICK
MATTHYSSENS, GASTON;WODAK, SHOSHANA;OVAX, WILHELMUS J.
TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
STABILITY OF PROTEINS
NUMBER OF SEQUENCES: 22
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/398,706
FILING DATE: 25-AUG-1989
SEQ ID NO:9;
LENGTH: 335
5290690-9

Query Match 56.7%; Score 59; DB 6; Length 335;
Best Local Similarity 84.6%; Pred. No. 0.093;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VLVGVNGFGRRGR 16
|11:|||||
DB 3 VKVINGFGRRGR 15

RESULT 13
US-09-095-855-205
Sequence 205, Application US/09095855
Patent No. 6160093
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth

APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 205:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-095-855-205

Query Match 56.7%; Score 59; DB 4; Length 340;
Best Local Similarity 84.6%; Pred. No. 0.094;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VLVGNGFGRIGR 16
: |||||
Db 3 IRVGVNGFGRIGR 15

RESULT 14
US-08-464-073-33
Sequence 33, Application US/08464073
Patent No. 5864028
GENERAL INFORMATION:
APPLICANT: Sloud, Mouldy
TITLE OF INVENTION: TNF-ALPHA RIBOZYMES, TNF-ALPHA RIBOZYME BINDING
TITLE OF INVENTION: PROTEIN AND DEGRADATION RESISTANT mRNA DERIVATIVES
TITLE OF INVENTION: LINKED TO TNF-ALPHA RIBOZYMES
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper and Dunham, LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,073
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/971,058
FILING DATE: 03-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU/1993/00567
FILING DATE: 03-NOV-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-464-073-33

Query Match 52.9%; Score 55; DB 2; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0087;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 VLVGNGFGRIG 15
: |||||
Db 1 VKVGVNGFGRIG 12

RESULT 15
5245013-20
Patent No. 5245013
APPLICANT: Ulevitch, Richard; Tobias, Peter
TITLE OF INVENTION: ACUTE PHASE PROTEIN MODULATING ENDOTOXIC
ACTIVITY OF LIPOPOLYSACCHARIDES, ASSAY METHODS AND POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/367,454
FILING DATE: 01-JUN-1989
APPLICATION NUMBER: 6,710
FILING DATE: 30-DEC-1986
APPLICATION NUMBER: 728,833
FILING DATE: 30-APR-1985
SEQ ID NO: 20:
LENGTH: 27
5245013-20

Query Match 50.0%; Score 52; DB 6; Length 27;
Best Local Similarity 66.7%; Pred. No. 0.062;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 6 VGVNGFGRIGRI 17
: :
Db 2 IAINGFGRIGRL 13

Search completed: June 13, 2001, 14:27:09
Job time: 630 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:25:50 ; Search time 150.28 Seconds
(without alignments)
7.608 Million cell updates/sec

Title: PCT-US01-05825A-32

Perfect score: 97
Sequence: 1 EVGLVESGGGLVQGRSLRL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /Geneseq_0401: *
2: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT: *
3: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT: *
4: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT: *
5: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT: *
6: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT: *
7: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT: *
8: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT: *
9: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT: *
10: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT: *
11: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT: *
12: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT: *
13: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT: *
14: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT: *
15: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT: *
16: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT: *
17: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT: *
18: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT: *
19: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT: *
20: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT: *
21: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT: *
22: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT: *
23: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	98	21	B40072
2	97	100.0	99	21	V50960
3	97	100.0	117	15	R52053
4	97	100.0	119	16	R66337
5	97	100.0	120	20	V43255
6	97	100.0	120	21	V50963
7	97	100.0	120	21	V50968
8	97	100.0	121	18	W27569
9	97	100.0	123	16	R66303
10	97	100.0	149	21	V64737
11	97	100.0	191	21	B53647

12	97	100.0	248	13	R20059	Recombinant sc3D6
13	97	100.0	475	13	R20057	Heavy chain of 3D6
14	97	100.0	483	22	B36207	Human Immune syste
15	94	96.9	30	17	R97323	Humanised monoclon
16	94	96.9	30	17	R97324	Humanised monoclon
17	94	96.9	110	18	W06207	Rat antibody heavy
18	94	96.9	115	18	W15522	Anti-TGF beta-2 sc
19	94	96.9	117	15	R32120	Heavy chain of rat
20	94	96.9	117	15	R32063	Heavy chain variab
21	94	96.9	117	16	R77305	Variable heavy cha
22	94	96.9	118	16	W06208	Xenograft antibody
23	94	96.9	119	16	W11919	Humanised MAb SK48
24	94	96.9	119	16	R82962	Humanised antibody
25	94	96.9	119	16	R82982	Humanised anti-VH
26	94	96.9	119	16	R82964	35193HUVHA humani
27	94	96.9	119	16	R82965	35193HUVHAS humani
28	94	96.9	119	16	R82966	35193HUVHAS human
29	94	96.9	120	15	R54930	Fc receptor human
30	94	96.9	120	20	V13948	VH chain of anti-C
31	94	96.9	122	21	V96065	Human anti-DAF ant
32	94	96.9	123	21	V71935	Heavy chain variab
33	94	96.9	140	18	W06205	Xenograft antibody
34	94	96.9	254	17	R94557	Humanised antibody
35	94	96.9	300	18	W11503	Humanised anti-Fc
36	94	96.9	300	20	W73218	H22-HRG fusion pro
37	94	96.9	352	20	V06272	Anti-Fc gamma rece
38	93	95.9	100	17	R96106	VH3-15 autoantibod
39	93	95.9	100	17	R95861	Anti-TGF beta-2 sc
40	93	95.9	117	18	W15523	LM609 antibody hea
41	93	95.9	117	19	W76003	VH domain CDR of a
42	92	94.8	120	20	V43254	Human 5' Esp relat
43	91	93.8	84	21	V64925	Human 5' Esp relat
44	91	93.8	95	21	B40127	Anti-hiL12 antibod
45	91	93.8	98	14	R34283	Human TNF binding

ALIGNMENTS

RESULT 1	
B40072	B40072 standard; Protein: 98 AA.
XX	
AC	B40072:
XX	
DT	05-FEB-2001 (first entry)
XX	
DE	Anti-hiL12 antibody H chain V region amino acid sequence SEQ ID 598.
XX	
KW	Human: neutralising antibody; Interleukin-12; IL-12; antiinflammatory;
KW	complementarily determining region; CDR; antitumoretic; antiarthritic;
KW	antisclerotic; neuroprotective; antipsoriatic; antiaschemic; cardiatic;
KW	antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
KW	multiple sclerosis; rheumatoid arthritis.
XX	
OS	Homo sapiens.
XX	
PN	W0200056772-A1.
XX	
PD	28-SEP-2000.
XX	
PF	24-MAR-2000; 2000WO-US07946.
XX	
PR	25-MAR-1999; 99US-0126603.
XX	
PA	(BADI) BASF AG.
XX	(GENW) GENETICS INST INC.
PI	Salfeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
PI	Kaymakcalan Z, Labkovsky B, Sakorafas P, Friedrich S, Wyles A;
PI	Veidman GM, Venturini A, Warne NW, Widom A, Eivlin JG, Duncan AR;
XX	Derbyshire EJ, Carmen S, Smith S, Hollett TL, Du Fou SL;

DR WPI; 2000-638250/61.
 XX
 PT New human antibody specific for human interleukin-12 (IL-12) used to
 PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
 PT disease and multiple sclerosis.
 XX
 PS Claim 75; Page 121; 377pp; English.
 XX
 CC This invention relates to a new human antibody specific for human
 CC interleukin-12 (IL-12). The invention also includes antigen binding
 CC portions that bind to IL-12. Sequences B39485-B39516 represent human
 CC anti-IL-12 antibody heavy and light chain complementarity determining
 CC region (CDR) amino acid sequences, and also includes variable region
 CC amino acid sequences. Other variable region amino acid sequences are
 CC given in B39517-B39560 and B40068-B40149. Sequences B39561-B39771
 CC represent anti-IL-12 CDR3 related amino acid sequences. B39772-B40063
 CC represent other CDR sequences. Light chain CDR3 consensus sequences are
 CC given in B40064-B40067. Primers used in the identification and
 CC construction of the antibodies of the invention are given in
 CC C61062-C61071. The antibody of the invention is a neutralising antibody
 CC and has antirheumatic, antiarthritic, antisclerotic, antiinflammatory;
 CC neuroprotective, antipsoriatic, antiasthmatic; cardiac; antiparasitic;
 CC antibacterial and immunosuppressive activity. The antibodies or
 CC antigen-binding fragments are useful in the treatment of disorders
 CC associated with detrimental release of human IL-12, especially Crohn's
 CC disease, multiple sclerosis and rheumatoid arthritis. They can also be
 CC used in the manufacture of a pharmaceutical composition to treat human
 CC IL-12 disorders.
 CC
 SQ Sequence 98 AA;
 XX
 SQ
 Query Match 100.0%; Score 97; DB 21; Length 98;
 Best Local Similarity 100.0%; Pred. No. 6.3e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EVLVESGGGLVOPGRSLRL 20
 ||||||||||||||||
 Db 1 evqlvesggglvqprslrl 20
 RESULT 2
 Y50960 standard; Protein; 99 AA.
 XX
 AC Y50960;
 XX
 DT 23-MAR-2000 (first entry)
 XX
 DE Human FVIII antibody A3-C1 scFv heavy chain protein DP-31.
 XX
 KW Human; heavy chain; antibody; factor VIII; hemostatic;
 KW hemophilia A; scFv; A3-C1.
 XX
 OS Homo sapiens.
 OS
 PN WO955680-A2.
 PN
 PD 18-NOV-1999.
 PD
 PF 07-MAY-1999; 99WO-NL00285.
 PF
 PR 08-MAY-1998; 98EP-0201543.
 PR
 PA (SANQ-) STICHTING SANQUIN BLOEDVOORZIENING.
 PA
 PI Voorberg JJ, Van Den Brink EN, Turenhout EAM;
 PI
 DR WPI; 2000-053102/04.
 DR
 PT New polynucleotide, polypeptide and antibody useful for diagnosing the
 PT presence of neutralizing antibodies against factor VIII and for
 PT treatment of hemophilia A patients with these antibodies -

XX
 PS Example 8; Fig 9A; 61pp; English.
 XX
 CC This invention describes a novel polynucleotide (1) (and complements and
 CC hybridizable polynucleotides) comprising a contiguous nucleotide sequence
 CC coding for a human antibody with factor VIII specificity which has
 CC hemostatic activity. (1) is useful as a primer or probe for detecting the
 CC presence of inhibitory antibodies directed against factor VIII. The
 CC polypeptides of the invention and the antibodies generated from them
 CC are useful in compositions for neutralizing factor VIII inhibiting
 CC antibodies in hemophilia A patients. This sequence represents the human
 CC factor VIII antibody A3-C1 specific scFv protein DP-31 which is used
 CC in the method of the invention.
 CC
 SQ Sequence 99 AA;
 XX
 SQ
 Query Match 100.0%; Score 97; DB 21; Length 99;
 Best Local Similarity 100.0%; Pred. No. 6.4e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EVLVESGGGLVOPGRSLRL 20
 ||||||||||||||||
 Db 1 evqlvesggglvqprslrl 20
 RESULT 3
 R52053 standard; Protein; 117 AA.
 XX
 AC R52053;
 XX
 DT 27-SEP-1996 (first entry)
 XX
 DE Heavy chain variable region of murine antibody IdFp.
 XX
 KW antibody; humanised; murine; human; heavy chain; light; variable;
 KW framework region; complementarity determining region; reshaping;
 KW modelling; surface residue; modify.
 XX
 OS Mus sp.
 OS
 PN EP592106-A1.
 PN
 PD 13-APR-1994.
 PD
 PF 07-SEP-1993; 93EP-0307051.
 PF
 PR 09-SEP-1992; 92US-0942245.
 PR
 PA (PEDE/) PEDERSEN J T.
 PA (IMMUT-) IMMUNOGEN INC.
 PA
 PI Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;
 PI
 DR WPI; 1994-120230/15.
 DR
 PT

Location/Qualifiers
 1..30 framework_region_1
 /label="FR 1"
 /note="FR 1"
 31..35
 /label="complementarity_determining_region_1
 /note="CDR 1"
 36..49
 /label="FR_2
 50..59
 /label="CDR_2
 60..98
 /label="FR_3
 99..115
 /label="CDR_3

KM keloid formation; scar tissue formation; pyrexia; HUVEC;
 KM periodontal disease; obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.
 OS Homo sapiens.
 XX
 XX WO9729131-A1.
 PN 14-AUG-1997.
 PD 10-FEB-1997; 97WO-0502219.
 PF 25-NOV-1996; 96US-0031476.
 PR 09-FEB-1996; 96US-0599226.
 XX
 XX (BADI) BASF AG.
 XX
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B;
 PI Manovich JA, McGuiness BT, Roberts AJ, Sakorats P;
 PI Seifeld JG, Schoenhaut D, Vaughan TJ, White M, Milton AJ;
 DR WPI: 1997-415302/38.
 DR N-PSDB: T88404.
 XX
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 XX
 PS Claim 16; Page 76; 102pp; English.
 XX
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain variable region.
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephritic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC
 SQ Sequence 121 AA:

Query Match 100.0%; Score 97; DB 18; Length 121;
 Best Local Similarity 100.0%; Pred. No. 8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVQLVESGGGLVQPGSRRL 20
 ||||||||||||||||
 DB 1 evqlvesggglvqpgsrsl 20

RESULT 9
 R66303 standard; Protein: 123 AA.
 ID R66303 standard; Protein: 123 AA.
 XX
 AC R66303;
 XX
 DT 02-AUG-1995 (first entry)
 XX
 DE Human immunoglobulin variable heavy chain #9.
 XX
 XX Primer: PCR; amplify; human; immunoglobulin; variable; heavy chain;

KM cosmid; placenta; vector; pJB81; E.coli; mammalian.
 XX
 OS Homo sapiens.
 XX
 XX WO9426895-A.
 PN 24-NOV-1994.
 PD 10-MAY-1993; 93WO-JP00603.
 PF 10-MAY-1993; 93WO-JP00603.
 PR 10-MAY-1993; 93WO-JP00603.
 XX
 XX (NISR) JAPAN TOBACCO INC.
 XX
 XX Honjo T, Matsuda F;
 PI WPI: 1995-006791/01.
 DR N-PSDB: Q78947.
 DR
 XX
 PT DNA fragment comprising human immunoglobulin Vh genes - for the
 PT production of human immunoglobulin in mammalian hosts
 XX
 PS Claim 18; Page 41-42; 130pp; Japanese.
 XX
 CC Protein sequences (R66295-51) are novel human immunoglobulin heavy chain
 CC sequences encoded by novel isolated genes. The genes (Q78939-79002) were
 CC isolated and cloned from a series of cosmid constructs: Y202; Y103; Y21;
 CC Y6; Y24; 3-31; M84; M18 and M131, by PCR amplification using primers
 CC Q78917-38. The genes are subdivided into 5 families of Vh genes. The
 CC fragments cover a region of 800 kb. The DNA fragments were isolated from
 CC high molecular weight DNA from human placenta. The DNA was partially
 CC digested with RsaI restriction enzyme. The fragments were separated by
 CC gel electrophoresis and 35-45 kb fractions were collected. The fragments
 CC were ligated with ClaI-digested cosmid vector pJB81. The ligation
 CC products were in vitro packaged and infected into E.coli 490A. The
 CC fragments were then subcloned by colony hybridisation. The Vh genes and
 CC the DNA fragments encoding them are useful in producing human
 CC immunoglobulin in mammalian hosts.
 CC
 SQ Sequence 123 AA:

Query Match 100.0%; Score 97; DB 16; Length 123;
 Best Local Similarity 100.0%; Pred. No. 8.1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVQLVESGGGLVQPGSRRL 20
 ||||||||||||||||
 DB 20 evqlvesggglvqpgsrsl 39

RESULT 10
 Y64737
 ID Y64737 standard; Protein: 149 AA.
 XX
 AC Y64737;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human 5' EST related polypeptide SEQ ID NO:898.
 XX
 DE Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
 KW gene therapy; chromosome mapping; upstream regulatory sequence;
 KW forensic; location; development; protein synthesis; stability;
 KW regulation; identification.
 XX
 XX OS Homo sapiens.
 XX
 XX PN WO9953051-A2.
 XX
 XX PD 21-OCT-1999.
 XX
 PF 09-APR-1999; 99WO-IB00712.

```

XX 09-APR-1998; 98US-0057719.
PR 28-APR-1998; 98US-0069047.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
DR WPI; 2000-038446/03.
DR N-PSDB; 242351.
XX
PT Novel secreted protein 5' expressed sequence tag sequences used in
XX diagnostic, forensic, gene therapy, and chromosome mapping procedures
PS Claim 3; Page 627; 837pp; English.
XX
CC 242265 to 243075 represent novel 5' expressed sequence tag (EST)
CC sequences, corresponding to human secreted proteins. Y64651 to Y65438
CC represent the EST-related proteins corresponding to 242265 to 243052.
CC The 5' ESTs can be used for producing secreted human gene products.
CC They can be used to identify and isolate 5' untranslated regions (UTRs)
CC and upstream regulatory regions which control the location, development
CC stage, rate, and quantity of protein synthesis, as well as stability of
CC mRNA. The ESTs are also useful as probes for chromosome mapping, and to
CC obtain full length cDNA clones. The ESTs can also be used in forensic
CC procedures to identify individuals, or in diagnostic procedures to
CC identify individuals having genetic diseases resulting from abnormal
CC gene expression. The products may also be used in gene therapy protocols.
CC The nucleic acids encoding signal peptides can be used for directing
CC extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.
CC The proteins encoded by the EST sequences may be useful in treating a
CC variety of human conditions. Secreted proteins have therapeutic value,
CC and the identification of new secreted proteins is valuable. 242249 to
CC 242264 and Y64644 to Y64650 represent sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 149 AA;

```

```

Query Match 100.0%; Score 97; DB 21; Length 149;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EVOLVESGGGLVQPGSRRL 20
   |||
Db 20 evglvesggglvqpgsrslrl 39

```

```

RESULT 11
ID B53647 standard; Protein; 191 AA.
XX
AC B53647;
XX
DT 09-MAR-2001 (first entry)
XX
DE Human colon cancer antigen protein sequence SEQ ID NO:1187.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW identification; cytostatic; cardioactive; neuroprotective; vulnerary;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder.
XX
OS Homo sapiens.
XX
PN WO200055351-A1.
XX
PD 21-SEP-2000.
XX

```

```

PF 08-MAR-2000; 2000WO-US05883.
XX
XX 12-MAR-1999; 99US-0124270.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
DR WPI; 2000-587534/55.
DR N-PSDB; C98404.
XX
PT Colon cancer associated gene sequences, referred to as colon cancer
XX antigens, useful for the treatment, prevention, and diagnosis of colon
XX disorders such as colon cancer -
PS Claim 11; Page 1767-1768; 2104pp; English.
XX
CC C97991 to C98763 encode the human colon cancer associated proteins,
CC called human colon cancer antigens, given in B53234 to B54006. The human
CC colon cancer antigens can have cytostatic, cardioactive, muscular,
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins
CC may also be used to prevent diseases such as neural disorders, immune
CC system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, wounds, renal disorders, infectious
CC diseases, and cardiovascular disorders. C98764 to C98772 and B54007
CC represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 191 AA;

```

```

Query Match 100.0%; Score 97; DB 21; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EVOLVESGGGLVQPGSRRL 20
   |||
Db 40 evglvesggglvqpgsrslrl 59

```

```

RESULT 12
ID R20059 standard; Protein; 248 AA.
XX
AC R20059;
XX
DT 25-MAR-1992 (first entry)
XX
DE Recombinant sc3d6 anti-HIV gp160 antibody.
XX
KW Plasmid pUC3d6LC; pUC3d6HC; human immunodeficiency virus; AIDS;
KW complementarity determining region.
XX
OS Homo sapiens.
XX
FH Key
FT Location/Qualifiers
FT 2..31
FT /label= Framework_1_heavy-chain
FT 32..36
FT /label= CDR_1_heavy-chain
FT 37..50
FT /label= Framework_2_heavy-chain
FT 51..67
FT /label= CDR_2_heavy-chain
FT 68..99
FT /label= Framework_3_heavy-chain
FT 100..116
FT Region

```



```

FT Region /label= CDR_3_heavy_chain
FT 117..127
FT /label= Framework_4_heavy_chain
FT 128..142
FT Region /label= linker
FT 143..165
FT Region /label= Framework_1_light_chain
FT 166..176
FT /label= CDR_1_light_chain
FT 177..191
FT Region /label= Framework_2_light_chain
FT 192..203
FT /label= CDR_2_light_chain
FT 204..235
FT /label= Framework_3_light_chain
FT 236..242
FT /label= CDR_3_light_chain
FT 243..253
FT Region /label= Framework_4_light_chain
FT
FT
XX WO9118983-A.
XX 12-DEC-1991.
XX
XX 28-MAY-1991; 91WO-1000067.
XX
XX 29-MAY-1990; 90AT-0001178.
XX
XX (JUNG/) JUNGBAUER A.
XX
XX Feigenhauer M, Himmler G, Kohl J, Steindl F;
XX N-PSDB: Q20068.
XX
XX Recombinant protein which binds to complex viral antigen and
XX HIV-1 - contains variable region of antibody derived from 3D6
XX cell line, used for detecting HIV-1 antigen
XX
XX Claim 4; Page 31; 52pp; German.
XX
XX The cell line 3D6 (87110301; Porton Down) produces a Mab of the
XX IgG1/kappa type that reacts specifically with HIV-1 gp41 and also
XX weekly cross-reacts with HIV-1 gp120. Gene construct sc3D6 was
XX engineered using the variable region coding regions of the heavy
XX CC and light chains of antibody 3D6, joined by a linker. The
XX CC recombinant protein binds to HIV gp160.
XX See also Q20066 and Q20067.
XX
XX SO Sequence 248 AA;

```

```

Query Match 100.0%; Score 97; DB 13; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 EVQLVESGGGLVQPGSRURL 20
    |||
DB 2 evqlvesggglvqpgrsrll 21
    |||

```

```

RESULT 13
R20057
ID R20057 standard; Protein; 475 AA.
XX
XX R20057;
XX
XX 25-MAR-1992 (first entry)
XX
XX Heavy chain of 3D6 anti-HIV antibody.
XX
XX Plasmid pUC3D6HC; human immunodeficiency virus; AIDS;
XX complementarity determining region.
XX
XX KW

```

```

XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..19
XX /label= signal
XX 20..49
XX /label= Framework_1
XX 50..54
XX /label= CDR-1
XX 55..68
XX /label= Framework_2
XX 69..85
XX /label= CDR_2
XX 86..117
XX /label= Framework_3
XX 118..134
XX /label= CDR_3
XX 135..145
XX /label= Framework_4
XX 146..475
XX /label= Constant_region
XX
XX WO9118983-A.
XX 12-DEC-1991.
XX
XX 28-MAY-1991; 91WO-1000067.
XX
XX 29-MAY-1990; 90AT-0001178.
XX
XX (JUNG/) JUNGBAUER A.
XX
XX Feigenhauer M, Himmler G, Kohl J, Steindl F;
XX N-PSDB: Q20066.
XX
XX Recombinant protein which binds to complex viral antigen and
XX HIV-1 - contains variable region of antibody derived from 3D6
XX cell line, used for detecting HIV-1 antigen
XX
XX Claim 2; Page 24; 52pp; German.
XX
XX The variable region of the heavy chain is used in a recombinant
XX CC protein with the variable region from the kappa light chain of 3D6,
XX CC the two V regions being joined by a linker. The recombinant protein
XX CC binds to HIV gp160.
XX See also Q20067 and Q20068.
XX
XX SO Sequence 475 AA;

```

```

Query Match 100.0%; Score 97; DB 13; Length 475;
Best Local Similarity 100.0%; Pred. No. 3.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 EVQLVESGGGLVQPGSRURL 20
    |||
DB 20 evqlvesggglvqpgrsrll 39
    |||

```

```

RESULT 14
B36207
ID B36207 standard; protein; 483 AA.
XX
XX B36207;
XX
XX 15-FEB-2001 (first entry)
XX
XX Human immune system associated protein HISAP-5.
XX
XX Human immune system associated protein; HISAP-5; immune disorder;
XX
XX KW

```

KW Infection; autoimmune disease; cancer.

OS Homo sapiens.

PN US6135941-A.

XX 24-OCT-2000.

PF 27-MAR-1998; 98US-0049672.

PR 27-MAR-1998; 98US-0049672.

PA (INCY-) INCYTE PHARM INC.

PI Tang YF, Yue H, Lal P, Corley NC, Guegler KJ, Baughn MR;

PI Hillman JL, Au-Young J;

DR WPI: 2001-030926/04.

DR N-PSDB; C66523.

PT New human immune system associated proteins (HISAP) and polynucleotides

PT encoding the HISAP, useful for diagnosing, treating or preventing

PS Claim 1; Column 55-58; 54pp; English.

CC The present invention provides the coding and protein sequences for a
CC number of human immune system associated proteins (HISAPs). These can be
CC used in the diagnosis and treatment of various autoimmune disorders,
CC infections and cell proliferation diseases. The diseases include AIDS,
CC adult respiratory distress syndrome, anaemia, asthma, atherosclerosis,
CC Crohn's disease, irritable bowel syndrome, multiple sclerosis, myasthenia
CC gravis, osteoarthritis, rheumatoid arthritis, scleroderma, systemic lupus
CC erythematosus, arteriosclerosis, cirrhosis and cancer.

Sequence 483 AA;

Query Match 100.0%; Score 97; DB 22; Length 483;

Best Local Similarity 100.0%; Pred. No. 3.5e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGSRSLRL 20

DB 20 evqlvesggglvqpgsrslrl 39

RESULT 15

R97323 R97323 standard; peptide; 30 AA.

AC R97323;

DT 15-OCT-1996 (first entry)

DE Humanised monoclonal antibody heavy chain framework region.

XX Monoclonal antibody; humanised; mouse; framework region; FR; CDR;

KW complementary determining region; anti-carcinoembryonic antigen;

KW CEA; diagnosis; imaging; therapy; immune response.

OS Homo sapiens.

XX WO9611013-A1.

PD 18-APR-1996.

PF 28-SEP-1995; 95WO-US11964.

PR 05-OCT-1994; 94US-0318157.

PA (IMMU-) IMMUNOMEDICS INC.

PI Armour KL, Hansen HJ;

DR WPI: 1996-209653/21.

PT New humanised anti-CEA monoclonal antibody - having engrafted murine

PT CDRs, used for diagnosis, imaging and therapy of CEA-producing

PS Claim 7; Page 39; 62pp; English.

CC New humanised monoclonal antibodies (MAbs) comprising the
CC complementary determining regions (CDRs) of a parental murine class
CC III anti-carcinoembryonic (CEA) Mab engrafted to the framework
CC regions (FRs) of a heterologous antibody which can be derived from
CC any species including human, retain the anti-CEA binding specificity
CC of the parental murine Mab but are less immunogenic in a human
CC subject than the parental Mab. The humanised antibodies can be used
CC in diagnosis, imaging and therapy of CEA-producing cancers and
CC patients receiving the humanised antibodies and conjugates show
CC improved therapeutic results, decreased immune responses and
CC decreased immune-mediated adverse effects compared to the parent
CC antibody. This sequence corresponds to the first framework region
CC of the heavy chain of the humanised Mab. See R97313-97333.

Sequence 30 AA;

Query Match 96.9%; Score 94; DB 17; Length 30;

Best Local Similarity 95.0%; Pred. No. 5.2e-08;

Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGSRSLRL 20

DB 1 evqlvesggglvqpgsrslrl 20

Search completed: June 13, 2001, 14:25:51
Job time: 670 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:23:13 ; Search time 87.97 Seconds

(Without alignments)
15.624 Million cell updates/sec

Title: PCT-US01-05825a-32

Sequence: 1 EVQLVESGGGLVQPGSRSLRL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	98	2	S26927 Ig heavy chain V r
2	97	100.0	100	2	S69896 Ig heavy chain V r
3	97	100.0	121	2	S31118 Ig heavy chain - h
4	97	100.0	121	2	S31104 Ig heavy chain (su
5	97	100.0	123	2	S30532 Ig heavy chain V r
6	97	100.0	128	2	S31585 Ig heavy chain V r
7	97	100.0	145	2	S11239 Ig heavy chain V r
8	94	96.9	97	2	S44115 Ig heavy chain V r
9	94	96.9	114	2	S46390 Ig heavy chain V r
10	94	96.9	123	2	PC4281 Ig heavy chain V r
11	91	93.8	94	2	anti-SS-A/Ro 60k p
12	91	93.8	98	2	PL0120 Ig heavy chain V-I
13	91	93.8	98	2	PL0116 Ig heavy chain V-I
14	91	93.8	113	2	S38450 Ig heavy chain - h
15	91	93.8	113	2	S57410 Ig heavy chain V-J
16	91	93.8	115	2	S57445 Ig heavy chain V-J
17	91	93.8	118	2	S31116 Ig heavy chain - h
18	91	93.8	119	2	F36005 Ig heavy chain V r
19	91	93.8	119	2	S31111 Ig heavy chain - h
20	91	93.8	120	2	S31112 Ig heavy chain - h
21	91	93.8	121	2	G36005 Ig heavy chain V r
22	91	93.8	122	2	E36005 Ig heavy chain V r
23	91	93.8	122	2	S31117 Ig heavy chain - h
24	91	93.8	122	2	S31119 Ig heavy chain - h
25	91	93.8	126	1	GIHUKL Ig heavy chain V-I
26	91	93.8	126	1	S48797 Ig heavy chain V r
27	91	93.8	130	2	S31601 Ig heavy chain V r
28	91	93.8	130	2	PL0098 Ig heavy chain pre

30	91	93.8	132	2	S31603 Ig heavy chain V r
31	91	93.8	133	2	A49028 Ig heavy chain V-I
32	91	93.8	133	2	S31590 Ig heavy chain V r
33	91	93.8	134	2	S31679 Ig heavy chain V r
34	91	93.8	137	2	S31701 Ig heavy chain V r
35	91	93.8	139	2	S31674 Ig heavy chain V r
36	90	92.8	97	2	S26890 Ig heavy chain V r
37	90	92.8	97	2	S46462 Ig heavy chain V r
38	90	92.8	98	2	PL0121 Ig heavy chain V-I
39	90	92.8	98	2	PL0123 Ig heavy chain V-I
40	90	92.8	98	2	S26896 Ig heavy chain V r
41	90	92.8	98	2	S29545 Ig heavy chain V r
42	90	92.8	98	2	S26932 Ig heavy chain V r
43	90	92.8	98	2	S26891 Ig heavy chain V r
44	90	92.8	98	2	S26894 Ig heavy chain V r
45	90	92.8	98	2	S26933 Ig heavy chain V r

ALIGNMENTS

RESULT 1

S26927 Ig heavy chain V region (DP-31) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

C:Accession: S26927

R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.

J. Mol. Biol. 227, 776-798, 1992

A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o

A:Reference number: S26885; MUID:93021117

A:Accession: S26927

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-98 <TOM>

A:Cross-references: EMBL:Z12333; NID:G32885; PIDN:CAA78203.1; PID:G32886

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 97; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 4.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVQLVESGGGLVQPGSRSLRL 20

DB 1 EVQLVESGGGLVQPGSRSLRL 20

RESULT 2

S69896 Ig heavy chain V region (clone RFK15H), rheumatoid factor - human

C:Species: Homo sapiens (man)

C>Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 23-Jul-1999

C:Accession: S69896

R:Randen, I.; Pascual, V.; Victor, K.; Thompson, K.M.; Forre, O.; Capra, D.J.; Natvig

Eur. J. Immunol. 23, 1220-1225, 1993

A:Title: Synovial IgG rheumatoid factors show evidence of an antigen-driven immune re

A:Reference number: S69896; MUID:93272805

A:Accession: S69896

A:Status: preliminary; translation not shown

A:Molecule type: mRNA

A:Residues: 1-100 <RAN>

A:Cross-references: EMBL:X73605; NID:G509797; PIDN:CAA51998.1; PID:G509798

C:Superfamily: immunoglobulin V region; immunoglobulin homology

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 97; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 4.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGRSRL 20
|||||
DB 1 EVOLVESGGGLVQPGRSRL 20

RESULT 3

S3118
Ig heavy chain - human
C:Species: Homo sapiens (man)
C:Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 17-Mar-1999
C:Accession: S3118
R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurman
Eur. J. Immunol. 22, 247-251, 1992
A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third complement
A:Reference number: S31104; WUID:92111633
A:Accession: S3118
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-121 <RAA>
A:Cross-references: EMBL:X62969
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 97; DB 2; Length 121;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGRSRL 20
|||||
DB 1 EVOLVESGGGLVQPGRSRL 20

RESULT 4

S31104
Ig heavy chain (subclass IgM) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 23-Jul-1999
C:Accession: S31104
R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurman
Eur. J. Immunol. 22, 247-251, 1992
A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third complement
A:Reference number: S31104; WUID:92111633
A:Accession: S31104
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-121 <RAA>
A:Cross-references: EMBL:X63080; NID:932646; PIDN:CAA44802.1; PID:932647
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 97; DB 2; Length 121;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGRSRL 20
|||||
DB 1 EVOLVESGGGLVQPGRSRL 20

RESULT 5

S30532
Ig heavy chain V region - human
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Aug-1996
C:Accession: S30532
R:Marlette, X.
submitted to the EMBL Data Library, October 1992

A:Reference number: S30520

A:Accession: S30532
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-123 <MAR>
A:Cross-references: EMBL:Z48318
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 97; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGRSRL 20
|||||
DB 1 EVOLVESGGGLVQPGRSRL 20

RESULT 6

S31595
Ig heavy chain V region - human
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S31595
R:Cuisinier, A.M.; Gauthier, L.; Boubli, L.; Fougereau, M.; Tonnelie, C.
submitted to the EMBL Data Library, June 1992
A:Description: Mechanisms that generate human immunoglobulin diversity operate from t
A:Reference number: S31585
A:Accession: S31595
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-128 <CUU>
A:Cross-references: EMBL:Z14171; NID:931007; PIDN:CAA78540.1; PID:931008
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:23-106/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 97; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 5.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGRSRL 20
|||||
DB 9 EVOLVESGGGLVQPGRSRL 28

RESULT 7

S11239
Ig heavy chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S11239
R:Felgenhauer, M.; Kohl, J.; Rueker, F.
Nucleic Acids Res. 18, 4927, 1990
A:Title: Nucleotide sequences of the cDNAs encoding the V-regions of H- and L-chains
A:Reference number: S11239; WUID:90370490
A:Accession: S11239
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-145 <FEL>
A:Cross-references: EMBL:X53613; NID:923865; PIDN:CAA37675.1; PID:9762936
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 97; DB 2; Length 145;
Best Local Similarity 100.0%; Pred. No. 6.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGRSRLR 20
| | | | | | | | | | | | | | | | | |
DB 20 EVOLVESGGGLVQPGRSRLR 39

RESULT 8
S44115
Ig heavy chain V region - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Jul-1999
C:Accession: S44115
R:Hawkins, R.E.; Zhu, D.; Ovecka, M.; Winter, G.; Hamblin, T.J.; Stevenson, F.K.
submitted to the EMBL data library, March 1994
A:Description: Idiotype vaccination against human B-cell lymphoma: rescue of variable r
A:Reference number: S44105
A:Accession: S44115
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-97 <RAW>
A:Cross-References: EMBL:231384; NID:9472969; PIDN:CAA83259.1; PID:9940526
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match 96.9%; Score 94; DB 2; Length 97;
Best Local Similarity 95.0%; Pred. No. 1.3e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGRSRLR 20
| | | | | | | | | | | | | | | | | |
DB 1 EVOLVESGGGVQPGRSRLR 20

RESULT 9
S46390
Ig heavy chain V region - human
C:Species: Homo sapiens (man)
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 20-Jun-2000
C:Accession: S46390
R:Figini, M.; Marks, J.D.; Winter, G.; Griffiths, A.D.
J. Mol. Biol. 239, 68-78, 1994
A:Title: In vitro assembly of repertoires of antibody chains on the surface of phage by
A:Reference number: S46390; MUID:94254092
A:Accession: S46390
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-114 <FIG>
A:Cross-References: EMBL:231686; NID:9509782; PIDN:CAA83491.1; PID:91335143
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 96.9%; Score 94; DB 2; Length 114;
Best Local Similarity 95.0%; Pred. No. 1.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGRSRLR 20
| | | | | | | | | | | | | | | | | |
DB 1 EVOLVESGGGVQPGRSRLR 20

RESULT 10
PC4281
anti-SS-A/Ro 60k peptide heavy chain E-56 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 21-Jan-2000
C:Accession: PC4281
R:Suzuki, H.; Takemura, H.; Suzuki, M.; Sekine, Y.; Kashiwagi, H.
Biochem. Biophys. Res. Commun. 232, 101-106, 1997
A:Title: Molecular cloning of anti-SS-A/Ro 60-kDa peptide fab fragments from infiltratin
A:Reference number: PC4279; MUID:97236289
A:Accession: PC4281

A:Molecule type: protein
A:Residues: 1-123 <SU2>
C:Comment: This antibody is commonly found in systemic autoimmune diseases such as Sj
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 96.9%; Score 94; DB 2; Length 123;
Best Local Similarity 95.0%; Pred. No. 1.6e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGRSRLR 20
| | | | | | | | | | | | | | | | | |
DB 1 EVOLVESGGGLVQPGRSRLR 20

RESULT 11
PL0120
Ig heavy chain V-III region (TD-Vo) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 16-Aug-1996
C:Accession: PL0120
R:Bird, J.; Gallil, N.; Link, M.; Stites, D.; Sklar, J.
J. Exp. Med. 168, 229-245, 1988
A:Title: Continuing rearrangement but absence of somatic hypermutation in Immunoglobu
A:Reference number: PL0116; MUID:88286083
A:Accession: PL0120
A:Molecule type: mRNA
A:Residues: 1-94 <BIR>
A:Experimental source: B cells from patient TD with acute lymphoblastic leukemia, ALL
A:Note: the sequence shows the V region (TD-Vo) from a nonproductive DNA rearrangemen
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: acute lymphoblastic leukemia; heterotetramer; Immunoglobulin
F:31-35/Region: complementarity-determining 1
F:49-65/Region: complementarity-determining 2

Query Match 93.8%; Score 91; DB 2; Length 94;
Best Local Similarity 90.0%; Pred. No. 3.7e-07;
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGRSRLR 20
: | | | | | | | | | | | | | | | | | |
DB 1 QVOLVESGGGVQPGRSRLR 20

RESULT 12
PL0116
Ig heavy chain V-III region (AW-Vx) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 23-Jul-1999
C:Accession: PL0116; S26892
R:Bird, J.; Gallil, N.; Link, M.; Stites, D.; Sklar, J.
J. Exp. Med. 168, 229-245, 1988
A:Title: Continuing rearrangement but absence of somatic hypermutation in Immunoglobu
A:Reference number: PL0116; MUID:88286083
A:Accession: PL0116
A:Molecule type: mRNA
A:Residues: 1-98 <BIR>
A:Experimental source: B cells from patient AW with acute lymphoblastic leukemia, ALL
A:Note: the sequence shows the V region (AW-Vx) from a nonproductive DNA rearrangemen
R:Tominson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o
A:Reference number: S26885; MUID:93021117
A:Accession: S26892
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-References: EMBL:212349; NID:932918; PIDN:CAA78219.1; PID:932919
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: acute lymphoblastic leukemia; heterotetramer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <IMM>

F:31-35/Region: complementarity-determining 1
F:49-65/Region: complementarity-determining 2

Query Match 93.8%; Score 91; DB 2; Length 98;
Best Local Similarity 90.0%; Pred. No. 3.8e-07;
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGSRSLRL 20
DB 1 QVQLVESGGGVQPGSRSLRL 20

RESULT 13

S29546
Ig heavy chain V region (COS-8 / DP-46) - human (fragment)

C:Species: Homo sapiens (man)
C:Date: 07-Jan-1994 #sequence_revision 17-Nov-1995 #text_change 23-Jul-1999
C:Accession: S29546; S26888
R:Tomlinson, M.; Walter, G.; Cook, G.P.; Winter, G.
submitted to the EMBL Data Library, October 1992
A:Reference number: S29543
A:Accession: S29546

A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z17394; NID:932843; PIDN:CAA78997.1; PID:932844
A:Note: designated COS-8
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline VH sequences reveals about fifty groups of V
A:Reference number: S26885; MUID:93021117
A:Accession: S26888

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TO2>
A:Cross-references: EMBL:Z12346; NID:932912; PIDN:CAA78216.1; PID:932913
A:Note: designated DP-46
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 93.8%; Score 91; DB 2; Length 98;
Best Local Similarity 90.0%; Pred. No. 3.8e-07;
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGSRSLRL 20
DB 1 QVQLVESGGGVQPGSRSLRL 20

RESULT 14

S38490
Ig heavy chain - human (fragment)

C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S38490
R:Marks, J.D.; Owehand, W.H.; Bye, J.M.; Finnern, R.; Gorlick, B.D.; Voak, D.; Thorpe, S.
submitted to the EMBL Data Library, June 1993
A:Description: Human antibody fragments specific for human blood group antigens from a F
A:Reference number: S38488
A:Accession: S38490
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <MAR>
A:Cross-references: EMBL:Z23030; NID:9414027; PIDN:CAA80565.1; PID:9414028
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 93.8%; Score 91; DB 2; Length 113;
Best Local Similarity 90.0%; Pred. No. 4.4e-07;

Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGSRSLRL 20
DB 1 QVQLVESGGGVQPGSRSLRL 20

RESULT 15

S57410
Ig heavy chain V-J region - human (fragment)

C:Species: Homo sapiens (man)
C:Date: 10-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 23-Jul-1999
C:Accession: S57410
R:Paterson, G.; Wilson, G.; Kennedy, P.G.E.; Willison, H.J.
submitted to the EMBL Data Library, June 1995
A:Description: Analysis of anti-GM1 ganglioside IgM antibodies cloned from motor neur
A:Reference number: S57408
A:Accession: S57410
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-113 <PAT>
A:Cross-references: EMBL:X87893; NID:9871363; PIDN:CAA61144.1; PID:9871364
A:Genetics: 99/2
A:Insertions: 99/2
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 93.8%; Score 91; DB 2; Length 113;
Best Local Similarity 90.0%; Pred. No. 4.4e-07;
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGSRSLRL 20
DB 1 QVQLVESGGGVQPGSRSLRL 20

Search completed: June 13, 2001, 14:23:13
Job time: 744 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:30:38 ; Search time 51.57 Seconds
(without alignments)
13.285 Million cell updates/sec

Title: PCT-US01-05825A-32
Perfect score: 97
Sequence: 1 EVOLVESGGLVQGRSLRL 20

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	93.8	126	1 HV3K_HUMAN	P01772 homo sapien
2	90	92.8	116	1 HV3T_HUMAN	P01781 homo sapien
3	90	92.8	119	1 HV3P_HUMAN	P01777 homo sapien
4	90	92.8	120	1 HV3E_HUMAN	P01766 homo sapien
5	90	92.8	120	1 HV3U_HUMAN	P01782 homo sapien
6	89	91.8	116	1 HV05_CARAU	P19181 carassius a
7	89	91.8	118	1 HV3V_HUMAN	P80419 homo sapien
8	88	90.7	119	1 HV3I_HUMAN	P01770 homo sapien
9	87	89.7	115	1 HV3D_HUMAN	P01765 homo sapien
10	87	89.7	116	1 HV3R_HUMAN	P01779 homo sapien
11	87	89.7	117	1 HV3Q_HUMAN	P01776 homo sapien
12	87	89.7	117	1 HV3O_HUMAN	P01774 homo sapien
13	87	89.7	119	1 HV3M_HUMAN	P01776 homo sapien
14	86	88.7	115	1 HV3F_HUMAN	P01767 homo sapien
15	86	88.7	117	1 HV17_MOUSE	P01786 mus musculu
16	86	88.7	122	1 HV20_MOUSE	P01789 mus musculu
17	86	88.7	122	1 HV21_MOUSE	P01790 mus musculu
18	86	88.7	122	1 HV3G_HUMAN	P01768 homo sapien
19	86	88.7	123	1 HV18_MOUSE	P01787 mus musculu
20	86	88.7	123	1 HV19_MOUSE	P01788 mus musculu
21	86	88.7	123	1 HV22_MOUSE	P01791 mus musculu
22	86	88.7	123	1 HV23_MOUSE	P01792 mus musculu
23	86	88.7	123	1 HV24_MOUSE	P01793 mus musculu
24	86	88.7	123	1 HV25_MOUSE	P01794 mus musculu
25	86	88.7	122	1 HV3H_HUMAN	P01769 homo sapien
26	83	85.6	98	1 HV57_MOUSE	P18528 mus musculu
27	83	85.6	115	1 HV3S_HUMAN	P01780 homo sapien
28	83	85.6	117	1 HV54_MOUSE	P18526 mus musculu
29	83	85.6	117	1 HV55_MOUSE	P18526 mus musculu
30	83	85.6	144	1 HV26_MOUSE	P01795 mus musculu
31	82	84.5	113	1 HV34_MOUSE	P01803 mus musculu
32	82	84.5	119	1 HV3N_HUMAN	P01775 homo sapien
33	81	83.5	114	1 HV3B_HUMAN	P01763 homo sapien

34	81	83.5	121	1 HV3I_HUMAN	P01771 homo sapien
35	80	82.5	116	1 HV36_MOUSE	P01806 mus musculu
36	80	82.5	117	1 HV41_MOUSE	P01811 mus musculu
37	80	82.5	119	1 HV37_MOUSE	P01807 mus musculu
38	80	82.5	119	1 HV38_MOUSE	P01808 mus musculu
39	80	82.5	119	1 HV40_MOUSE	P01810 mus musculu
40	80	82.5	122	1 HV3A_HUMAN	P01762 homo sapien
41	79	81.4	97	1 HV56_MOUSE	P18527 mus musculu
42	79	81.4	114	1 HV01_CANFA	P01784 canis famil
43	78	80.4	117	1 HV58_MOUSE	P18529 mus musculu
44	78	80.4	136	1 HV16_MOUSE	P01783 mus musculu
45	77	79.4	119	1 HV3L_HUMAN	P01773 homo sapien

ALIGNMENTS

RESULT	ID	STANDARD	PRT	126 AA.
1	HV3K_HUMAN			
AC	P01772;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	IG HEAVY CHAIN V-III REGION KOL.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
NCBI	Taxid=9606;			
RN	[1]			
RP	SEQUENCE AND DISULFIDE BONDS.			
RA	MEDLINE-83289131; PubMed-6884994;			
RA	Schmidt W.E.; Jung H.-D.; Palm W.; Hilschmann N.;			
RT	"Three-dimensional structure determination of antihodies. Primary			
RT	structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.;"			
RL	Hope-Seayler's Z. Physiol. Chem. 364:713-747(1983).			
RN	[2]			
RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).			
RX	MEDLINE-81072293; PubMed-7441755;			
RA	Marquart M.; Deisenhofer J.; Huber R.; Palm W.;			
RT	"Crystallographic refinement and atomic models of the intact			
RT	immunoglobulin molecule KOL and its antigen-binding fragment at 3.0 A			
RT	and 1.0-A resolution.;"			
RL	J. Mol. Biol. 141:365-391(1980).			
DR	PIR; A02055; G1HUKL.			
DR	PDB; 2EB4; 12-JUL-89.			
DR	PDB; 2IG2; 12-JUL-89.			
DR	InterPro: IPR003006; "			
DR	Pfam: PF00047; 19; 1;			
KW	Immunoglobulin V region; 3D-structure.			
FT	MOD_RES	1		
FT	DISULFID	22	96	
FT	DISULFID	105	110	
FT	STRAND	3	7	
FT	STRAND	11	12	
FT	TURN	14	15	
FT	TURN	18	25	
FT	STRAND	18	25	
FT	HELIX	29	31	
FT	STRAND	34	39	
FT	TURN	41	42	
FT	TURN	46	51	
FT	STRAND	53	54	
FT	TURN	58	60	
FT	STRAND	62	67	
FT	TURN	68	73	
FT	STRAND	74	77	
FT	TURN	78	83	
FT	STRAND	88	90	
FT	HELIX	92	99	
FT	STRAND	106	106	
FT	TURN	107	108	
FT	STRAND	109	109	
FT	STRAND	113	113	

FT STRAND 120 124
 FT NON-TER 126 126
 SQ SEQUENCE 126 AA; 13718 MW; EAD71B52B16F8776 CRC64;

Query Match
 Best Local Similarity 93.8%; Score 91; DB 1; Length 126;
 Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVOPGRSLRL 20
 : |||||
 DB 1 EVOLVESGGGLVOPGRSLRL 20

RESULT 2
 HV3P_HUMAN STANDARD; PRT; 116 AA.

AC P01781;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION GAL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.

RA MEDLINE=75059123; PubMed=4803843;
 RX Matanabe S., Barnikol H.U., Horn J., Bertram J., Hilschmann N.;
 RT "The primary structure of a monoclonal Igm-immunoglobulin
 RT (macroglobulin Gal.); II: the amino acid sequence of the H-chain (mu-
 RT type), subgroup H III. Architecture of the complete Igm-molecule."
 RN Hoppe-Seyler's Z. Physiol. Chem. 354:1505-1509(1973).
 RP REVISION TO THE COMPOSITION OF 28-33.

RA Hilschmann N.;
 RL Submitted (JUN-1975) to the PIR data bank.

CC -I- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM A WALDENSTROM'S

CC MACROGLOBULIN.

DR PIR; A02064; M3HUGL.

DR HSSP; P01772; 21G2.

DR InterPro; IPR003006; -

DR Pfam; PF00047; 1g; 1.

KW Immunoglobulin V region.

FT NON-TER 116 116

SQ SEQUENCE 116 AA; 12730 MW; 2C67CA9AAAAA1282 CRC64;

Query Match
 Best Local Similarity 92.8%; Score 90; DB 1; Length 116;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVOPGRSLRL 20
 : |||||
 DB 1 EVOLVESGGGLVOPGRSLRL 20

RESULT 3
 HV3P_HUMAN STANDARD; PRT; 119 AA.

AC P01777;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION TER.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.

MEMLINE=74142702; PubMed=4522793;

RA Capra J.D., Kehoe J.M.;
 RT "Variable region sequences of five human immunoglobulin heavy chains
 RT of the VH3 subgroup: definitive identification of four heavy chain
 RT hypervariable regions."
 RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
 CC PROTEIN.

DR PIR; A02060; G1HOTE.
 DR HSSP; P01772; 21G2.
 DR InterPro; IPR003006; -
 DR Pfam; PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT NON-TER 119 119
 SQ SEQUENCE 119 AA; 12802 MW; 7E24DC652C7290A9 CRC64;

Query Match
 Best Local Similarity 92.8%; Score 90; DB 1; Length 119;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVOPGRSLRL 20
 : |||||
 DB 1 EVOLVESGGGLVOPGRSLRL 20

RESULT 4
 HV3E_HUMAN STANDARD; PRT; 120 AA.

AC P01766;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION BRO.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.

RA MEDLINE=77117674; PubMed=65324;

RX Capra J.D., Hopper J.E.;

RT "Comparative studies on monotypic Igm lambda and Igg kappa from an

RT individual patient. III. The complete amino acid sequence of the VH

RL region of the Igm paraprotein."

CC Immunochimistry 13:995-999(1976).

CC -I- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM IGM ISOLATED FROM THE

CC SERUM OF A PATIENT WITH MALIGNANT LYMPHOMA OF THE WALDENSTROM

CC TYPE.

DR PIR; A02049; M3HUGB.

DR InterPro; IPR003006; -

DR Pfam; PF00047; 1g; 1.

KW Immunoglobulin V region.

FT NON-TER 120 120

SQ SEQUENCE 120 AA; 13227 MW; D3F0428F7C2E6410 CRC64;

Query Match
 Best Local Similarity 92.8%; Score 90; DB 1; Length 120;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVOPGRSLRL 20
 : |||||
 DB 1 EVOLVESGGGLVOPGRSLRL 20

RESULT 5
 HV3U_HUMAN STANDARD; PRT; 120 AA.

AC P01782;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION DOB.

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=80020921; Pubmed=114209;
RA Steiner L.A., Garcia Pardo A., Margolies M.N.;
RT Amino acid sequence of the heavy-chain variable region of the
RL crystallizable human myeloma protein Dob.";
RL Biochemistry 18:4068-4080(1979).
RN [2]
RP CRYSTALLIZATION.
RX MEDLINE=80020920; Pubmed=114208;
RA Steiner L.A., Lopes A.D.;
RT "The crystallizable human myeloma protein Dob has a hinge-region
RL deletion.";
RL Biochemistry 18:4054-4067(1979).
CC -I- MISCELLANEOUS: THIS GAMMA-1 MYELOMA PROTEIN HAS A DELETION IN THE
CC HINGE REGION. THERE ARE NO LIGHT-HEAVY OR INTER-HEAVY CHAIN
CC DISULFIDE BONDS.
DR PIR: A02065; GIHUBD.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13440 MW; 880DDE307C4B2627 CRC64;

Query Match
Best Local Similarity 92.8%; Score 90; DB 1; Length 120;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVPGPGRSLRL 20
DQ 1 EVOLVESGGGLVPGPGRSLRL 20

RESULT 6
HV05_CARAU STANDARD; PRT; 116 AA.
AC P19181;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION 5A PRECURSOR.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8814476; Pubmed=2125551;
RA Wilson M.R., Middleton D., Watt G.W.;
RT "Immunoglobulin heavy chain variable region gene evolution: structure
RT and family relationships of two genes and a pseudogene in a teleost
RT fish.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 116 IG HEAVY CHAIN V REGION 5A.
FT DOMAIN 20 49 FRAMEWORK 1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING 1.
FT DOMAIN 55 68 FRAMEWORK 2.
FT DOMAIN 69 84 COMPLEMENTARITY-DETERMINING 2.
FT DOMAIN 85 116 FRAMEWORK 3.
FT DISULFID 41 114 BY SIMILARITY.
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 12808 MW; 9C2279E2DF199B12 CRC64;

```

```

Query Match
Best Local Similarity 91.8%; Score 89; DB 1; Length 116;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVPGPGRSLRL 20
DQ 20 EVOLVESGGGLVPGPGRSLRL 39

RESULT 7
HV3V_HUMAN STANDARD; PRT; 118 AA.
ID HV3V_HUMAN
AC P80419;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION GAR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=95255298; Pubmed=7737190;
RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
RT "Characterization of the two unique human anti-flavin monoclonal
RT immunoglobulins.";
RL Eur. J. Biochem. 228:886-893(1995).
DR HSSP: P01810; 2FBJ.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 118 118
SQ SEQUENCE 118 AA; 13087 MW; 6C21D810ED1B6D1F CRC64;

Query Match
Best Local Similarity 91.8%; Score 89; DB 1; Length 118;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVPGPGRSLRL 20
DQ 1 EVOLVESGGGLVPGPGRSLRL 20

RESULT 8
HV3L_HUMAN STANDARD; PRT; 119 AA.
ID HV3L_HUMAN
AC P01770;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION NIE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77070269; Pubmed=826475;
RA Postling H., Hilschmann N.;
RT "The rule of antibody structure. The primary structure of a
RT monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The
RT chymotryptic peptides of the H-chain, alignment of the tryptic
RT peptides and discussion of the complete structure.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
RN [2]
RP DISULFIDE BOND.
RX MEDLINE=77070267; Pubmed=1002129;
RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
RT "Rule of antibody structure. The primary structure of a monoclonal

```

RT 19g1 immunoglobulin (myeloma protein Nie). I: Purification and
 RT characterization of the protein, the L- and H-chains, the
 RT cyanogen bromide cleavage products, and the disulfide bridges."
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
 CC PROTEIN.
 DR PIR: A02033; G1HUNT.
 DR HSSP: P01772; 21G2.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT MOD_RES 1 1
 FT DISULFID 22 96
 FT NON_TER 119 139
 SQ SEQUENCE 119 AA; 13242 MW; C96935A6E55E165B CRC64;

Query Match 90.7%; Score 88; DB 1; Length 119;
 Best Local Similarity 85.0%; Pred. No. 5.1e-07;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EVOLVESGGGLVOPGSRSLRL 20
 DB 1 EVOLVESGGGLVOPGSRSLRL 20

RESULT 9
 ID HV3D_HUMAN STANDARD; PRT; 115 AA.
 AC P01765:
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-II REGION TIL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=78005528; PubMed=409716;
 RA Wang A.-C., Wang I.Y., Fudenberg H.H.;
 RT "Immunoglobulin structure and genetics. Identify between variable
 RT regions of a mu and a gamma2 chain."
 RL J Biol. Chem. 252:7192-7199(1977).
 CC -I- MISCELLANEOUS: THE SEQUENCES OF THE V REGIONS OF THE HEAVY CHAINS
 CC OF IGM AND IGG2 ISOLATED FROM A SINGLE PATIENT WITH BICLONAL
 CC GAMMOPATHY ARE IDENTICAL. THEIR LIGHT CHAINS ARE APPARENTLY ALSO
 CC IDENTICAL.
 DR PIR: A02048; H3HUNT.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT NON_TER 115 115
 SQ SEQUENCE 115 AA; 12356 MW; 4DCC67D179F62326 CRC64;

Query Match 89.7%; Score 87; DB 1; Length 115;
 Best Local Similarity 90.0%; Pred. No. 7.1e-07;
 Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 EVOLVESGGGLVOPGSRSLRL 20
 DB 1 EVOLVESGGGLVOPGSRSLRL 20

RESULT 10
 ID HV3R_HUMAN STANDARD; PRT; 116 AA.
 AC P01779:
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE IG HEAVY CHAIN V-II REGION TUR.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=74142702; PubMed=4522793;
 RA Capra J.D., Kehoe J.M.;
 RT "Variable region sequences of five human immunoglobulin heavy chains
 RT of the VH3 subgroup: definitive identification of four heavy chain
 RT hypervariable regions."
 RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGA1 MYELOMA
 CC PROTEIN.
 DR PIR: A02062; A1HUNT.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT NON_TER 116 116
 SQ SEQUENCE 116 AA; 12431 MW; EB705F53A963F0C CRC64;

Query Match 89.7%; Score 87; DB 1; Length 116;
 Best Local Similarity 90.0%; Pred. No. 7.2e-07;
 Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 EVOLVESGGGLVOPGSRSLRL 20
 DB 1 EVOLVESGGGLVOPGSRSLRL 20

RESULT 11
 ID HV3C_HUMAN STANDARD; PRT; 117 AA.
 AC P01764:
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-II REGION VH26 PRECURSOR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81101090; PubMed=6450418;
 RA Mathysens G., Rabbits T.H.;
 RT "Structure and multiplicity of genes for the human immunoglobulin
 RT heavy chain variable region."
 RL Proc. Natl. Acad. Sci. U.S.A. 77:6561-6565(1980).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: J00236; AAA53516.1; -
 DR EMBL: M35415; AAA58735.1; -
 DR PIR: A02047; H3H026.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 117 IG HEAVY CHAIN V-II REGION VH26.
 FT NON_TER 117 117
 SQ SEQUENCE 117 AA; 12582 MW; E826733F1A3CB0F1 CRC64;

Query Match 89.7%; Score 87; DB 1; Length 117;

RT. (MOPC 47 A) with a 100-residue deletion."
RL J. Biol. Chem. 254:11418-11430(1979).
CC -1- MISCELLANEOUS: THIS ALPHA CHAIN WAS ISOLATED FROM A MYELOMA
CC PROTEIN THAT CONTAINS ONE LIGHT AND ONE HEAVY CHAIN PER MOLECULE,
CC LINKED BY A DISULFIDE BOND. IN CONTRAST, NORMAL MOUSE IGA
CC MOLECULES CONTAIN TWO LIGHT AND TWO HEAVY CHAINS AND LACK A
CC LIGHT-HEAVY CHAIN DISULFIDE BOND.
DR PIR; A02069; AIMS47.
DR HSSP; P01789; 2MCP.
DR InterPro; IPR003006; .
DR Pfam; PF00047; 1g; 1.
KW Immunoglobulin V region.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12975 MW; 0C74BE8BB154BDF4 CRC64;

Query Match 88.7%; Score 86; DB 1; Length 117;
Best Local Similarity 90.0%; Pred. No. 1e-06;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGGGIYOPGSRSLRL 20
11:|||||1111111111111111
Db 1 EVKLVESSGGGIYOPGSRSLRL 20

Search completed: June 13, 2001, 14:30:38
Job time: 528 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:29:47 ; Search time 150.43 Seconds
(without alignments)
15.583 Million cell updates/sec

Title: PCT-US01-05825A-32
Perfect score: 97
Sequence: 1 EVOLVESGGGLVQPGKRSURL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP-REMBL_15:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP-organellar:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP_unclassified:*
13: SP-vertebrate:*
14: SP-virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	96.9	122	4	09UL84
2	90	92.8	118	4	09UL91
3	90	92.8	118	4	09UL72
4	89	91.8	116	4	09UL93
5	87	89.7	113	4	09UL90
6	87	89.7	121	4	09UL71
7	86	88.7	131	4	09UL88
8	83	85.6	147	4	09UL509
9	79	81.4	95	4	09UL86
10	75	77.3	15	11	09UL16
11	72	74.2	437	11	09UL14
12	67	69.1	298	11	09ULF0
13	66	68.0	16	4	09UL53
14	62	63.9	124	6	09UL04
15	59	60.8	124	6	09UL06
16	56	57.7	119	4	09UL73
17	55	56.7	119	4	09UL94
18	55	56.7	124	4	09UL92
19	55	56.7	125	4	09UL95

20	54	55.7	117	11	09ULF0	09ULF0 mus musculu
21	54	55.7	117	11	09ULX9	09ULX9 mus musculu
22	54	55.7	122	4	09UL75	09UL75 homo sapien
23	53	54.6	150	4	09UL93	09UL93 homo sapien
24	51	52.6	117	11	09UL16	09UL16 mus musculu
25	50	51.5	157	4	09UL98	09UL98 homo sapien
26	49	50.5	150	4	09UL298	09UL298 homo sapien
27	47	48.5	89	2	09ULF0	09ULF0 thermus agu
28	46	47.4	449	2	09UL06	09UL06 streptomyc
29	45	46.4	336	5	09ULB05	09ULB05 streptomyc
30	45	46.4	500	10	023154	023154 arabidopsis
31	45	46.4	1695	4	09ULP3	09ULP3 homo sapien
32	45	46.4	2540	4	09UL02	09UL02 homo sapien
33	44	45.4	678	10	065709	065709 arabidopsis
34	44	45.4	1435	4	09ULB9	09ULB9 homo sapien
35	44	45.4	1550	4	09UL47	09UL47 homo sapien
36	44	45.4	1798	5	09ULB6	09ULB6 drosophila
37	43.5	44.8	342	4	09UL383	09UL383 homo sapien
38	43	44.3	168	2	09UL2V7	09UL2V7 streptomyc
39	43	44.3	292	2	09UL701	09UL701 streptomyc
40	43	44.3	233	10	09UL789	09UL789 arabidopsis
41	43	44.3	402	11	035444	035444 mus musculu
42	43	44.3	632	2	09ULF1	09ULF1 pseudomonas
43	43	44.3	860	14	073309	073309 human immun
44	43	44.3	860	14	073310	073310 human immun
45	43	44.3	860	14	073312	073312 human immun

ALIGNMENTS

RESULT 1
ID 09UL84
AC 09UL84
IC 09UL84
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035030; AAD56266.1; -;
DR HSSP; P01772; 2PBA.
DR INTERPRO; IPR003006; -;
DR PRAM; PF00047; 1g; 1.1
FT NON_TER
FT NON_TER
FT NON_TER
SQ SEQUENCE 122 AA; 13579 MW; 36054D41366545B8 CRC64;

Query Match 96.9%; Score 94; DB 4; Length 122;
Best Local Similarity 95.0%; Pred. NO. 5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVQPGKRSURL 20
DB 1 EVOLVESGGGLVQPGKRSURL 20
RESULT 2
ID 09UL91
AC 09UL91; PRELIMINARY; PRT; 118 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-JUN-2000 (TREMBLrel. 13, Last sequence update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035023; AAD56259.1; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 118 AA; 12843 MW; D0633949FZAC149D CRC64;

Query Match 92.8%; Score 90; DB 4; Length 118;
Best Local Similarity 95.0%; Pred. No. 2e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGSRSLR 20
Db 1 EVOLVESGGGLVQPGSRSLR 20

RESULT 3
O9UL72 PRELIMINARY; PRT; 118 AA.
AC O9UL72;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035042; AAD56278.1; -.
DR HSSP: P01772; 2F84.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 118 AA; 12872 MW; BAD1A5944B2D5CCA CRC64;

Query Match 92.8%; Score 90; DB 4; Length 118;
Best Local Similarity 95.0%; Pred. No. 2e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGSRSLR 20
Db 1 EVOLVESGGGLVQPGSRSLR 20

RESULT 4
O9UL93

ID O9UL93 PRELIMINARY; PRT; 116 AA.
AC O9UL93;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035021; AAD56257.1; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 116 AA; 12434 MW; ODA0348154DD6061 CRC64;

Query Match 91.8%; Score 89; DB 4; Length 116;
Best Local Similarity 94.7%; Pred. No. 2.7e-06;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VOLVESGGGLVQPGSRSLR 20
Db 1 VOLVESGGGLVQPGSRSLR 19

RESULT 5
O9UL90 PRELIMINARY; PRT; 113 AA.
AC O9UL90;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035024; AAD56260.1; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 113 AA; 12437 MW; ED57FDD19086D07F CRC64;

Query Match 89.7%; Score 87; DB 4; Length 113;
Best Local Similarity 90.0%; Pred. No. 5.4e-06;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGSRSLR 20
Db 1 EVOLVESGGGLVQPGSRSLR 20

RESULT 6

RESULT 10
 090V16 PRELIMINARY; PRT; 15 AA.
 AC 090V16;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE PROLACTIN-BINDING PROTEIN (FRAGMENT).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE.
 RC MEDLINE=95094032; PubMed=8000909;
 RA Cohen H., Cohen O., Gagnon J.;
 RT "Serum prolactin-binding protein (PRL-SP) of human and rat are
 RT identified as 19g.";
 RL C. R. Acad. Sci., III, Sci. Vie 317:293-298(1994).
 DR HSSP: P01789; IMCP.
 SQ SEQUENCE 15 AA; 1469 MW; 35ED2512FF3FA369 CRC64;

Query Match 77.3%; Score 75; DB 11; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVOPG 15
 Db 1 EVOLVESGGGLVOPG 15

RESULT 11
 09R1A4 PRELIMINARY; PRT; 437 AA.
 ID 09R1A4;
 AC 09R1A4;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GAMMA1 HEAVY CHAIN OF MAB7 (FRAGMENT).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
 RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
 RT antibody (Mab 7, its light and heavy chains) and construction of a
 RT single chain antibody (scFv)."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF152372; AD40243.1; -
 DR HSSP: P01842; 7FAB.
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; 19; 4.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN; 1.
 FT NON_TER 1
 FT 437
 SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D657C CRC64;

Query Match 74.2%; Score 72; DB 11; Length 437;
 Best Local Similarity 78.9%; Pred. No. 0.0046;
 Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 VOLVESGGGLVOPGRSLRL 20
 Db 1 VOLVESGGGLVOPGRSLRL 19

RESULT 12

090YFO PRELIMINARY; PRT; 298 AA.
 ID 090YFO;
 AC 090YFO;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE CN 8 SCFV.
 CN CN 8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE=SPLEEN.
 RA Shinozaki N., Demura T., Fukuda H.;
 RT "Isolation of a novel type of vascular cell wall-specific monoclonal
 RT antibody recognizing a cell polarity using a phage display subtraction
 RT method."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB036341; BAA8633.1; -
 DR HSSP: P01607; 1RET.
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; 19; 2.
 SQ SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;

Query Match 69.1%; Score 67; DB 11; Length 298;
 Best Local Similarity 65.0%; Pred. No. 0.018;
 Matches 13; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVOPGRSLRL 20
 Db 40 QVKLOQSGGLVOPGRSLRL 59

RESULT 13
 09UC53 PRELIMINARY; PRT; 16 AA.
 ID 09UC53;
 AC 09UC53;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE 77 KDA SPONTANEOUS RECURRENT ABORTION-ASSOCIATED HUMAN EMBRYONIC
 DE ANTIGEN/IGVHII HOMOLOG (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RC MEDLINE=96033130; PubMed=8582963;
 RA Shiraishi Y., Shiraishi Y., Yamamoto D., Hasegawa T., Kitamura W.,
 RA Miki S., Tanaka T., Suzuki T., Soma H.;
 RT "Diagnostic relevance of abortion-associated human embryonic antigen
 RT expressed on the cell surface of tumour promoter-treated Bloom
 RT syndrome cells."
 RL Hum. Reprod. 10:1694-1701(1995).
 SQ SEQUENCE 16 AA; 1626 MW; C9C5ED2512FF3FB9 CRC64;

Query Match 68.0%; Score 66; DB 4; Length 16;
 Best Local Similarity 93.3%; Pred. No. 0.001;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVOPG 15
 Db 1 EVOLVESGGGLVOPG 15

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:27:09 ; Search time 78.71 Seconds
(without alignments)
4.881 Million cell updates/sec

Title: PCT-US01-05825A-32

Sequence: 1 EVOLVESGGGLVQPGKSLRL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	100.0	117	1	US-07-942-245-24
2	97	100.0	118	3	US-08-545-809A-97
3	97	100.0	119	3	US-08-545-809A-131
4	97	100.0	120	1	US-08-478-039-08
5	97	100.0	120	1	US-08-478-039-08
6	97	100.0	120	1	US-08-478-039-08
7	97	100.0	121	3	US-08-599-226-2
8	97	100.0	125	1	US-08-478-039-99
9	97	100.0	125	1	US-08-478-039-99
10	97	100.0	483	4	US-09-049-672A-5
11	94	96.9	30	2	US-08-318-157B-30
12	94	96.9	30	2	US-08-318-157B-31
13	94	96.9	117	1	US-07-942-245-34
14	94	96.9	117	2	US-08-652-558-3
15	94	96.9	117	2	US-08-652-558-44
16	94	96.9	117	2	US-08-652-558-45
17	94	96.9	117	2	US-08-652-558-46
18	94	96.9	119	1	US-08-207-996-22
19	94	96.9	119	2	US-08-859-649-20
20	94	96.9	119	2	US-08-859-649-21
21	94	96.9	119	2	US-08-859-649-22
22	94	96.9	119	2	US-08-859-649-23
23	94	96.9	119	2	US-08-859-649-24
24	94	96.9	119	2	US-08-859-649-31
25	94	96.9	119	2	US-08-318-157B-12
26	94	96.9	119	2	US-08-318-157B-13
27	94	96.9	119	2	US-08-318-157B-14

28	94	96.9	119	2	US-08-318-157B-15	Sequence 15, Appl
29	94	96.9	119	2	US-08-318-157B-17	Sequence 17, Appl
30	94	96.9	119	2	US-08-318-157B-58	Sequence 58, Appl
31	94	96.9	119	2	US-08-760-840A-22	Sequence 22, Appl
32	94	96.9	119	4	US-09-266-119-22	Sequence 22, Appl
33	94	96.9	119	5	PCT-US94-07659-6	Sequence 6, Appl
34	94	96.9	121	3	US-08-599-226-10	Sequence 10, Appl
35	94	96.9	126	2	US-08-318-157B-7	Sequence 7, Appl
36	94	96.9	300	2	US-08-661-052-4	Sequence 4, Appl
37	94	96.9	489	5	PCT-US95-11405-35	Sequence 35, Appl
38	93	95.9	100	1	US-08-320-515B-3	Sequence 3, Appl
39	93	95.9	100	1	US-08-309-025-3	Sequence 3, Appl
40	92	94.8	120	2	US-08-958-201-8	Sequence 8, Appl
41	91	93.8	98	1	US-08-211-202-118	Sequence 118, App
42	91	93.8	116	1	US-08-211-202-141	Sequence 141, App
43	91	93.8	117	1	US-07-942-245-18	Sequence 18, App
44	91	93.8	117	1	US-07-942-245-36	Sequence 36, App
45	91	93.8	117	3	US-08-545-809A-115	Sequence 115, App

ALIGNMENTS

RESULT 1
US-07-942-245-24
Sequence 24, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUILD, Braydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sugrue, Mion, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-24

Query Match 100.0%; Score 97; DB 1; Length 117;

Best Local Similarity 100.0%; Pred. No. 5.4e-08; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVOLVESGGGLVQPGKSLRL 20
|||||
DB 1 EVOLVESGGGLVQPGKSLRL 20

RESULT 2
US-08-545-809A-97
; Sequence 97, Application US/08545809A
; Patent No. 6096878
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
; NUMBER OF SEQUENCES: 145
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 27-MAR-1996
; APPLICATION NUMBER: US/08/545,809A
; PRIORITY APPLICATION DATA:
; FILING DATE: 10-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 06501/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEEX: 200154
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 118 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-545-809A-97
Query Match 100.0%; Score 97; DB 3; Length 118;
Best Local Similarity 100.0%; Pred. No. 5.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVOLVESGGGLVOPGRSLRL 20
Db 20 EVOLVESGGGLVOPGRSLRL 39
RESULT 3
US-08-545-809A-131
; Sequence 131, Application US/08545809A
; Patent No. 6096878
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
; NUMBER OF SEQUENCES: 145
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEEX: 200154
INFORMATION FOR SEQ ID NO: 131:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-131
Query Match 100.0%; Score 97; DB 3; Length 119;
Best Local Similarity 100.0%; Pred. No. 5.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVOLVESGGGLVOPGRSLRL 20
Db 20 EVOLVESGGGLVOPGRSLRL 39
RESULT 4
US-08-478-039-98
; Sequence 98, Application US/08478039
; Patent No. 5681722
; GENERAL INFORMATION:
; APPLICANT: Newman, Roland A.
; APPLICANT: Hanna, Nabill
; TITLE OF INVENTION: Recombinant Antibodies for Human Therapy
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince St.
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,039
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/379,072
; FILING DATE: 25-JAN-1995
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/912,292
; FILING DATE: 10-JUL-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/856,281
; FILING DATE: 23-MAR-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/735,064

FLYING DATE: 25-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Teskin Esq., Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 012712-160
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
POSITION IN GENOME:
CHROMOSOME/SEGMENT: CHERRI
US-08-478-039-98

Query Match 100.0%; Score 97; DB 1; Length 120;
Best Local Similarity 100.0%; Pred. No. 5.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVOPGRSLRL 20
DB 1 EVOLVESGGGLVOPGRSLRL 20

RESULT 5
US-08-476-349A-98
Sequence 98, Application US/08476349A
Patent No. 3750105
GENERAL INFORMATION:
APPLICANT: Newman, Roland A.
APPLICANT: Hanna, Nabil W.
APPLICANT: Raab, Ronald W.
TITLE OF INVENTION: Recombinant Antibodies for Human Therapy
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: 699 Prince St.
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,349A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/379,072
FILING DATE: 25-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/912,292
FILING DATE: 10-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/856,281
FILING DATE: 23-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/735,064
FILING DATE: 25-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Teskin Esq., Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 012712-161
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620

TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
POSITION IN GENOME:
CHROMOSOME/SEGMENT: CHERRI
US-08-476-349A-98

Query Match 100.0%; Score 97; DB 1; Length 120;
Best Local Similarity 100.0%; Pred. No. 5.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVOPGRSLRL 20
DB 1 EVOLVESGGGLVOPGRSLRL 20

RESULT 6
US-08-958-201-10
Sequence 10, Application US/08958201
Patent No. 5977319
GENERAL INFORMATION:
APPLICANT: Pope, Anthony R.
APPLICANT: Pritchard, Kevin
APPLICANT: Williams, Andrew J.
APPLICANT: Johnson, Kevin S.
TITLE OF INVENTION: Specific binding members for estradiol;
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall O'Toole Gerstein Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.35 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/958,201
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/028,897
FILING DATE: 21-OCT-1996
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: 2D8
US-08-958-201-10

Query Match 100.0%; Score 97; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 5.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVOPGRSLRL 20
DB 1 EVOLVESGGGLVOPGRSLRL 20

RESULT 7
US-08-599-226-2

Sequence 2, Application US/08599226
Patent No. 6090382
GENERAL INFORMATION:
APPLICANT: Salfeld, Jochen G.
APPLICANT: Allen, Deborah J.
APPLICANT: Hoogenboom, Hendricus R.J.M.
APPLICANT: Kaymakalan, Zehra
APPLICANT: Labkovsky, Boris
APPLICANT: Mankovich, John A.
APPLICANT: McGuinness, Brian T.
APPLICANT: Roberts, Andrew J.
APPLICANT: Sakorafas, Paul
APPLICANT: Schoenhaut, David
APPLICANT: Vaughan, Tristan J.
APPLICANT: White, Michael
TITLE OF INVENTION: Human Antibodies that Bind Human TNFa
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,226
FILING DATE: 08-FEB-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A., Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-043
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 121 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
US-08-599-226-2

Query Match 100.0%; Score 97; DB 3; Length 121;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVQPGSRRL 20
DB 1 EVOLVESGGGLVQPGSRRL 20

RESULT 8
US-08-478-039-99
Sequence 99, Application US/08478039
Patent No. 5681722
GENERAL INFORMATION:
APPLICANT: Newman, Roland A.
APPLICANT: Hanna, Nabil
APPLICANT: Raab, Ronald W.
TITLE OF INVENTION: Recombinant Antibodies for Human Therapy
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: 699 Prince St.

CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,039
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/379,072
FILING DATE: 25-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/912,292
FILING DATE: 10-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/856,281
FILING DATE: 23-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/735,064
FILING DATE: 25-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Teskin Esq., Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 012712-160
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 125 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 18/2
US-08-478-039-99

Query Match 100.0%; Score 97; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 5.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVQPGSRRL 20
DB 1 EVOLVESGGGLVQPGSRRL 20

RESULT 9
US-08-476-349A-99
Sequence 99, Application US/08476349A
Patent No. 5750105
GENERAL INFORMATION:
APPLICANT: Newman, Roland A.
APPLICANT: Hanna, Nabil
APPLICANT: Raab, Ronald W.
TITLE OF INVENTION: Recombinant Antibodies for Human Therapy
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: 699 Prince St.
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,349A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/379,072
FILING DATE: 25-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/912,292
FILING DATE: 10-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/856,281
FILING DATE: 23-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/735,064
FILING DATE: 25-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Teskin Esq., Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 012712-161
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 125 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 18/2
US-08-476-349A-99

Query Match 100.0%; Score 97; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 5.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVOLVESGGGLVOPGRSLRL 20
Db 1 EVOLVESGGGLVOPGRSLRL 20

RESULT 10
US-09-049-672A-5
Sequence 5, Application US/09049672A
Patent No. 6135941
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Tang, Y. Tom
APPLICANT: Yue, Henry
APPLICANT: Au-Young, Janice
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: HUMAN IMMUNE SYSTEM ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/049,672A
FILING DATE: HERewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0497 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 483 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BMAROT03
CLONE: 1669829
US-09-049-672A-5

Query Match 100.0%; Score 97; DB 4; Length 483;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVOLVESGGGLVOPGRSLRL 20
Db 20 EVOLVESGGGLVOPGRSLRL 39

RESULT 11
US-08-318-157B-30
Sequence 30, Application US/08318157B
Patent No. 5874540
GENERAL INFORMATION:
APPLICANT: HANSEN, Hans J.
APPLICANT: ARMOUR, Kathryn L.
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,157B
FILING DATE: 05-OCT-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 18733/464
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-318-157B-30

Query Match 96.9%; Score 94; DB 2; Length 30;
Best Local Similarity 95.0%; Pred. No. 3.6e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGVOPGRSLRL 20
DB 1 EVOLVESGGGVOPGRSLRL 20

RESULT 12
US-08-318-157B-31
Sequence 31, Application US/08318157B
Patent No. 5874540

GENERAL INFORMATION:
APPLICANT: HANSEN, Hans J.
APPLICANT: ARMOUR, Kathryn L.
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
NUMBER OF SEQUENCES: 58.
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,157B
FILING DATE: 05-OCT-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 18/33/464
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-318-157B-31

Query Match 96.9%; Score 94; DB 2; Length 30;
Best Local Similarity 95.0%; Pred. No. 3.6e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGVOPGRSLRL 20
DB 1 EVOLVESGGGVOPGRSLRL 20

RESULT 13
US-07-942-245-34
Sequence 34, Application US/07942245
Patent No. 5639641

GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUILD, Braydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
TITLE OF INVENTION: ANTIBODIES
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-34

Query Match 96.9%; Score 94; DB 1; Length 117;
Best Local Similarity 95.0%; Pred. No. 1.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGVOPGRSLRL 20
DB 1 EVOLVESGGGVOPGRSLRL 20

RESULT 14
US-08-652-558-3
Sequence 3, Application US/08652558
Patent No. 5861155
GENERAL INFORMATION:
APPLICANT: LIN, AUGUSTINE YEE-THARN
TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER & WITCOFF
STREET: 75 STATE STREET, 23RD FLOOR
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,558
FILING DATE: JUNE 6, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB94/00387

FILED DATE: NOVEMBER 21, 1994
ATTORNEY/AGENT INFORMATION:
NAME: YANKWICH, LEON R.
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: 95,497-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-558-3

Query Match 96.9%; Score 94; DB 2; Length 117;
Best Local Similarity 95.0%; Pred. No. 1.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGRSRL 20
|||||
Db 1 EVOLVESGGGVQPGRSRL 20

RESULT 15
US-08-652-558-44
Sequence 44, Application US/08652558
Patent No. 5861155
GENERAL INFORMATION:
APPLICANT: LIN, AUGUSTINE YEE-THARN
TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER & WITCOFF
STREET: 75 STATE STREET, 23RD FLOOR
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,558
FILING DATE: JUNE 6, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB94/00387
FILING DATE: NOVEMBER 21, 1994
ATTORNEY/AGENT INFORMATION:
NAME: YANKWICH, LEON R.
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: 95,497-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-558-44

Query Match 96.9%; Score 94; DB 2; Length 117;
Best Local Similarity 95.0%; Pred. No. 1.5e-07;

Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPGRSRL 20
|||||
Db 1 EVOLVESGGGVQPGRSRL 20

Search completed: June 13, 2001, 14:27:09
Job time: 630 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:25:51 ; Search time 150.28 Seconds
(without alignments)
5.325 Million cell updates/sec

Title: PCT-US01-05825A-33

Perfect score: 79
Sequence: 1 FCLGRCLVODGFVT 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_0401:*

1: /SID66/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SID66/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SID66/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SID66/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SID66/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SID66/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SID66/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SID66/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SID66/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SID66/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SID66/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SID66/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SID66/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SID66/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SID66/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SID66/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SID66/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SID66/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SID66/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SID66/gcgdata/geneseq/geneseq/AA2000.DAT:*
21: /SID66/gcgdata/geneseq/geneseq/AA2001.DAT:*
22: /SID66/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	399	13 R20006	Zonula occludens t
2	45	57.0	8	21 Y79110	Peptide antagonist
3	45	57.0	751	21 G39904	Arabidopsis thalia
4	45	57.0	781	21 G39903	Arabidopsis thalia
5	45	57.0	837	21 G39902	Arabidopsis thalia
6	43	54.4	284	21 G24558	Arabidopsis thalia
7	43	54.4	742	16 R74094	Human zona pelluci
8	43	54.4	745	15 R55206	Human zona pelluci
9	43	54.4	745	20 Y42480	Human zona pelluci
10	43	54.4	745	20 W81817	Human ZPA protein
11	43	54.4	745	21 Y82215	Human zona pelluci

12	43	54.4	745	21 Y52689	Human oocyte zona
13	43	54.4	745	21 Y52180	Human zona pelluci
14	43	54.4	745	21 Y52984	Human zona pelluci
15	42.5	53.8	84	20 Y30424	Mature nematode ex
16	42.5	53.8	162	20 Y30437	Mature nematode ex
17	42.5	53.8	162	21 B15322	A. caninum nematod
18	42.5	53.8	181	17 R91712	AcAMP47. Ancylus
19	42.5	53.8	181	20 Y30410	Nematode extracted
20	42	53.2	8	21 Y79106	Peptide antagonist
21	41	51.9	13	19 W52566	Cyclic peptide of
22	41	51.9	200	20 Y34767	Chlamydia pneumon
23	41	51.9	713	15 R60101	Canine zona pelluc
24	41	51.9	715	15 R55198	Canine zona pelluc
25	41	51.9	715	20 Y42471	Canine zona pelluc
26	41	51.9	715	20 W81808	Canine ZPA protein
27	41	51.9	715	21 Y82206	Canine zona pelluc
28	41	51.9	715	21 Y52680	Canine oocyte zona
29	41	51.9	715	21 Y52171	Canine oocyte zona
30	41	51.9	715	21 Y52975	Canine zona pelluc
31	41	51.9	716	15 R55200	Canine zona pelluc
32	41	51.9	716	15 R60532	Canine zona pelluc
33	41	51.9	716	20 Y42473	Feline zona pelluc
34	41	51.9	716	20 W81810	Feline ZPA protein
35	41	51.9	716	21 Y82208	Feline zona pelluc
36	41	51.9	716	21 Y52682	Feline oocyte zona
37	41	51.9	716	21 Y52173	Feline zona pelluc
38	41	51.9	716	21 Y52977	Feline zona pelluc
39	40	50.6	713	11 R06998	Mouse ZP2 protein
40	38	48.1	8	21 Y79109	Peptide antagonist
41	38	48.1	8	21 Y79122	Peptide antagonist
42	38	48.1	36	21 B40159	Human secreted pro
43	38	48.1	903	21 B25110	Eucalyptus grandis
44	37	46.8	8	21 Y79114	Peptide antagonist
45	37	46.8	713	15 R55194	Porcine zona pellu

ALIGNMENTS

RESULT 1	
R20006	R20006 standard; Protein: 399 AA.
XX	
AC	R20006;
XX	
DT	31-MAR-1992 (first entry)
XX	
DE	Zonula occludens toxin.
XX	
KW	ZOR: cholera; vaccine; enterotoxin; diarrhoea.
XX	
OS	Vibrio cholerae.
XX	
PN	W09118979-A.
XX	
PD	12-DEC-1991.
XX	
PF	05-JUN-1991; 91WO-US03812.
XX	
PR	05-JUN-1990; 90US-0533315.
XX	
PA	(UYMA-) UNIV MARYLAND BALTI.
XX	
PI	Kaper JB, Baudry-Maurelli B, Fasano A;
XX	
DR	WPI: 1992-007465/01.
XX	
DR	N-PSDB; Q20185.
XX	
PT	New Vibrio cholerae strains - comprise restriction endonuclease
XX	
FT	fragment encoding toxin, used as vaccines against cholera
XX	
PS	Disclosure: Fig 18; 83pp; English.

CC The amino acid sequence is that of the zonula occludens toxin (ZOT).
CC It may be responsible for diarrhoea in some strains of cholera and
CC the ZOT gene or fragments of it are deleted from strains of *Vibrio*
CC *cholerae* (V.c.) to be used as vaccines. These V.c. strains have 100%
CC efficiency in protecting humans against subsequent infection with a
CC strain of a similar serotype and avoid undesirable side effects such
CC as diarrhoea, nausea and cramping. Cultures of these strains may be
CC used for prodn. of vaccines against cholera.
XX
SQ Sequence 399 AA;

Query Match 100.0%; Score 79; DB 13; Length 399;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Caps 0;
Db 288 fcigrlcvdgvfvt 301
1 FCIGRLCVDGVFVT 14
|||||
Y79110
Y79110 standard; Peptide: 8 AA.
XX
AC Y79110;
XX
DT 05-JUN-2000 (first entry)
XX
DE Peptide antagonist of zonulin.
XX
KM Zonulin; antagonist; zonula occludens toxin receptor;
KM blood-brain barrier; antiinflammatory; cerebroprotective;
KM neuroprotective; dermatologically; antitumor; antiviral;
KM antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
KM hypotensive; immunosuppressive; antiparasitic; vasotropic;
KM gastrointestinal inflammation; therapy.
XX
OS Synthetic.
XX
PN WO200007609-A1.
XX
PD 17-FEB-2000.
XX
PF 28-JUL-1999; 99WO-US16683.
XX
PR 03-AUG-1998; 98US-0127815.
XX
XX (UYMA-) UNIV MARYLAND BALTIMORE.
XX
PA Fasano A;
XX
PI WPI; 2000-205565/18.
XX
DR
XX
PT New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
XX
PS Claim 1; Page 42; 69pp; English.
XX
CC This present sequence is that of a peptide antagonist of zonulin
CC (Z), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor. yet do not
CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical synthesis or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for

CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.
XX
SQ Sequence 8 AA;

Query Match 57.0%; Score 45; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Caps 0;
Db 1 grlcvgdg 8
4 GRICVODG 11
|||||
G39904
G39904 standard; Protein: 751 AA.
XX
AC G39904;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49440.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
XX
OS
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134768.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139453.
PR 16-JUN-1999;	99US-0139457.
PR 17-JUN-1999;	99US-0138492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0138458.
PR 18-JUN-1999;	99US-0138459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 01-JUL-1999;	99US-0142154.
PR 02-JUL-1999;	99US-0142055.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143542.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.

PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151530.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
PR 21-OCT-1999;	99US-0160815.
PR 22-OCT-1999;	99US-0160980.
PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160989.
PR 25-OCT-1999;	99US-0161404.

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 57.0%; Score 45; DB 21; Length 751;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 FCIGRLCYODGF 12
:|:|:|:|:
Db 84 yclgrmcvcpf 95

RESULT 4
G39903
ID G39903 standard; Protein; 781 AA.

AC G39903;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 49439.

KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay: genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-030139.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 57.0%; Score 45; DB 21; Length 781;
Best Local Similarly 50.0%; Pred. No. 23;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 1 FCIGRLCYQDGF 12
Db 114 yclgmcvpepf 125
RESULT 5
ID G39902 standard; Protein; 837 AA.
XX AC G39902;
XX DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49438.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PE 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151303.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159332.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 57.0%; Score 45; DB 21; Length 837;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 FCIGRLCVQDGE 12
Db 170 yclgrmcvpepf 181

RESULT 6
G24558
ID G24558 standard; Protein; 284 AA.
XX
AC G24558;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28278.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126254.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

```

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 54.4%; Score 43; DB 21; Length 284;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FCIGRLC 7
 |||||
 Db 208 fcigrlc 214

RESULT 7
 R74094

```

ID R74094 standard; Protein; 742 AA.
XX
AC R74094;
XX
DT 04-JAN-1996 (first entry)
XX
DE Human zona pellucida-2 (HZP-2) protein.
XX
KW Human zona pellucida; HZP-2; protein; vaccine; antigen;
XX contraceptive; ovary.
XX
OS Homo sapiens.
XX
PN JP0709974-A.
XX
PD 18-APR-1995.
XX
PE 05-OCT-1993; 93JP-0249404.
XX
PR 05-OCT-1993; 93JP-0249404.
XX
PA (TOFU ) TONEN CORP.
XX
DR WPI; 1995-182067/24.
XX
DR N-PSDB; Q92254.
XX
PT A DNA coding human zona pellucida-2 protein - used as a vaccine
XX antigen
XX
PS Claim 1; Page 8-11; 11pp; Japanese.
XX
CC The human zona pellucida-2 (HZP-2) protein may be used as a vaccine
XX antigen. It can be artificially synthesised using recombinant
XX techniques. Partial sequences of the HZP-2 gene were subcloned, the
XX sequences and primers used in cloning and subcloning are given in
XX Q92255-69.
XX
SQ Sequence 742 AA;

Query Match 54.4%; Score 43; DB 16; Length 742;
Best Local Similarity 70.0%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 GRUCVODGFV 13
  |||||
  Db 369 gelctgqdgfm 378

RESULT 8
R55206
ID R55206 standard; Protein; 745 AA.
XX
AC R55206;
XX
DT 01-FEB-1995 (first entry)
XX
DE Human zona pellucida ZPA protein.
XX
KW Human; zona pellucida; ZPA; immunoreception.
XX
OS Homo sapiens.
XX
PN Key
XX Key Location/Qualifiers
XX Protein 1..745
XX /label= human_ZPA

OY 1 FCIGRLC 7
  |||||
  Db 208 fcigrlc 214

RESULT 7
R74094

```

```

PR 09-NOV-1992; 92US-0973341.
PR 29-JAN-1993; 93US-0012990.
PA (ZONA-) ZONAGEN INC.
XX
XX Harris JD, Hsu KT, Podolski JS;
XX WPI: 1994-183156/22.
DR N-PSDB; Q65616.
XX
XX Use of zona pellucida proteins and antibodies - for inducing
XX reproducible transient infertility or permanent sterility in
XX female mammals
XX
XX Claim 40; Page 126-128; 154pp; English.
PS
XX A commercially available human genomic DNA library consisting of
XX 9-23kb inserts of DNA from placenta tissue of a male caucasian
XX cloned into the Lambda fix (RM)II vector was screened with
XX probes encoding porcine ZPA and ZPB proteins. Positive clones
XX were analysed further by Southern hybridisation using the porcine
XX probes and restriction analysis. Clones encoding human ZPA and ZPB
XX proteins were identified (see Q65616 and Q65617, respectively).
XX R55206 is the deduced amino acid sequence for human ZPA protein.
XX
XX Sequence 745 AA:
SQ
Query Match 54.4%; Score 43; DB 15; Length 745;
Best Local Similarity 70.0%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 4 GRLCVQDGFV 13
DB 369 gelctqdgfm 378
RESULT 9
Y42480
ID Y42480 standard; Protein; 745 AA.
XX
XX Y42480;
XX
XX 15-FEB-2000 (first entry)
XX
XX Human zona pellucida ZPA protein.
XX
XX Contraceptive; porcine; pig; zona pellucida; mammal; lapine; canine;
XX dog; rabbit; feline; cat; bovine; cow; human; cynomolgus monkey; ZPA;
XX ZPB; ZPC; Infertility; Veterinary.
XX
XX Homo sapiens.
XX
XX US5981228-A.
XX
XX 09-NOV-1999.
XX
XX 07-JUN-1995; 95US-0484596.
XX
XX 09-NOV-1993; 93US-0149223.
XX
XX 09-NOV-1992; 92US-0973341.
XX
XX 29-JAN-1993; 93US-0012990.
XX
XX (ZONA-) ZONAGEN INC.
XX
XX Podolski JS, Hsu KT, Harris JD;
XX
XX WPI: 1999-633318/54.
XX
XX N-PSDB; Z22737.
XX
XX Mammalian zona pellucida proteins used to induce transient or permanent
XX Infertility -

```

```

PS Example 11; Column 127-132; 84pp; English.
XX
XX This sequence represents the human zona pellucida ZPA protein. The
XX invention relates to the isolation of novel nucleotide sequences encoding
XX zona pellucida proteins from mammalian, especially porcine, lapine,
XX canine, feline, bovine, human or cynomolgus monkey sources. The zona
XX pellucida proteins (ZPA, ZPB, and ZPC) are used in the induction of
XX transient or permanent infertility. At present the method is used in
XX veterinary applications to induce transient or permanent infertility
XX in porcine, lapine, canine, feline, bovine, and cynomolgus monkeys.
XX
XX Sequence 745 AA:
SQ
Query Match 54.4%; Score 43; DB 20; Length 745;
Best Local Similarity 70.0%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 4 GRLCVQDGFV 13
DB 369 gelctqdgfm 378
RESULT 10
W81817
ID W81817 standard; Protein; 745 AA.
XX
XX W81817;
XX
XX 29-JAN-1999 (first entry)
XX
XX Human ZPA protein.
XX
XX ZPA; zona pellucida; infertility; sterility; immunocontraceptive;
XX vaccine; human.
XX
XX Homo sapiens.
XX
XX US5837497-A.
XX
XX 17-NOV-1998.
XX
XX 07-JUN-1995; 95US-0484993.
XX
XX 09-NOV-1993; 93US-0149223.
XX
XX 09-NOV-1992; 92US-0973341.
XX
XX 29-JAN-1993; 93US-0012990.
XX
XX 07-JUN-1995; 95US-0484993.
XX
XX (ZONA-) ZONAGEN INC.
XX
XX Harris JD;
XX
XX WPI: 1999-023447/02.
XX
XX N-PSDB; V64814.
XX
XX Isolated zona pellucida DNA from different mammals - used to develop
XX products which can be used for vaccination to induce transient
XX infertility or permanent sterility in female mammals
XX
XX Example 11; Column 133-138; 84pp; English.
XX
XX This sequence represents a human ZPA protein isolated from zona
XX pellucida. This protein can be used in a method for specifically
XX inducing transient infertility or permanent sterility in a host
XX animal by selective vaccination with specific zona pellucida proteins
XX or immunocontraceptively active fragments.
XX
XX Sequence 745 AA:
SQ
Query Match 54.4%; Score 43; DB 20; Length 745;
Best Local Similarity 70.0%; Pred. No. 47;

```

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GRICVODGFV 13
 | | | | |
 Db 369 gelctqdgfm 378

RESULT 11
 Y82215
 ID Y82215 standard; Protein; 745 AA.

AC Y82215;
 DT 12-JUN-2000 (first entry)
 XX
 DE Human zona pellucida ZPA SEQ ID NO:43.

KW Zona pellucida; ZPA; ZPB; ZPC; Infertility; permanent sterility;
 KM vaccine; immunononreceptive; contraceptive; immunostimulatory.

OS Homo sapiens.

PN US6027727-A.

PD 22-FEB-2000.

PF 09-NOV-1993; 93US-0149223.

PR 09-NOV-1992; 92US-0973341.

PR 29-JAN-1993; 93US-0012990.

PA (ZONA-) ZONAGEN INC.

PI Podolski JS, Hsu KT, Harris JD;

DR WPI; 2000-269144/23.

DR N-PSDB; 295674.

PT Inducing reproducible transient or permanent infertility in a mammal
 PT comprises administration of homologous and/or heterologous mammalian
 PT species zona pellucida proteins -

XX Example 11; Column 135-138; 85pp; English.

XX The present invention describes a method for inducing reproducible
 CC transient infertility in a female mammal, including humans, by selective
 CC administration of homologous and/or heterologous mammalian species
 CC zona pellucida (ZP) proteins or immunononreceptively active fragments.
 CC Also described are: (1) a method for inducing transient infertility in a
 CC mammal by administering directly an antibody directed to a ZPB or an
 CC immunologically active fragment selected from feline, bovine,
 CC cynomolgus monkey or human ZPB polypeptides; (2) an isolated, purified
 CC recombinant feline, bovine, cynomolgus monkey or human ZPB polypeptide
 CC or immunopolypeptide active fragment; and (3) a fusion protein comprising
 CC a ZPB polypeptide which is conjugated with a compound selected from
 CC keyhole limpet haemocyanin, muramyl dipeptide, histidine-tag, beta-gal,
 CC or palmitic acid where the fusion protein remains effective to stimulate
 CC production of antibodies that recognise a ZPB polypeptide. The method is
 CC useful for providing transient or permanent infertility or sterility in
 CC humans and animals. The present sequence represents human ZPA, which is
 CC used in the exemplification of the present invention.

XX Sequence 745 AA;

Query Match 54.4%; Score 43; DB 21; Length 745;
 Best Local Similarity 70.0%; Pred. No. 47;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GRICVODGFV 13
 | | | | |
 Db 369 gelctqdgfm 378

RESULT 12

Y52689
 ID Y52689 standard; Protein; 745 AA.

AC Y52689;

DT 14-MAR-2000 (first entry)

DE Human oocyte zona pellucida protein ZPA.

KW Zona pellucida; ZPA; ZPB; ZPC; purified; mammalian; glycoprotein;
 KM target; immunononreception; vaccine; antibody; transient; infertility;
 KM controllable; predictable; permanent; sterility; side effect.

OS Homo sapiens.

PN US6001599-A.

PD 14-DEC-1999.

PF 02-JUN-1995; 95US-0458731.

PR 09-NOV-1993; 93US-0149223.

PR 09-NOV-1992; 92US-0973341.

PR 29-JAN-1993; 93US-0012990.

PA (ZONA-) ZONAGEN INC.

PI Podolski JS, Hsu KT, Harris JD;

DR WPI; 2000-061880/05.

PT Isolated DNA encoding mammalian zona pellucida proteins useful for
 PT inducing transient sterility -

XX Example 11; Columns 133-136; 84pp; English.

XX This sequence represents a mammalian zona pellucida protein, ZPA.
 CC The zona pellucida is a complex matrix surrounding the mammalian
 CC oocyte, formed of glycoproteins secreted by ovarian cells. The
 CC invention relates to zona pellucida proteins ZPA, ZPB and ZPC, which
 CC are useful as targets for immunononreceptive vaccines. Polynucleotides
 CC encoding these proteins can be used for recombinant protein production,
 CC and as probes in hybridisation methods for the isolation of
 CC polynucleotides encoding zona pellucida proteins from other mammalian
 CC species. Administration of zona pellucida proteins to a host animal
 CC results in a specific immunononreceptive effect. Administration of
 CC purified ZPA and/or ZPB induces transient infertility in females.
 CC Fertility can be maintained in an on or off condition in a controllable/
 CC predictable fashion. Administration of ZPC induces permanent sterility.
 CC Use of a purified zona pellucida protein rather than a complex mixture
 CC of zona pellucida proteins reduces the potential for unwanted side-
 CC effects which may be harmful.

XX Sequence 745 AA;

Query Match 54.4%; Score 43; DB 21; Length 745;
 Best Local Similarity 70.0%; Pred. No. 47;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GRICVODGFV 13
 | | | | |
 Db 369 gelctqdgfm 378

RESULT 13

Y52180
 ID Y52180 standard; Protein; 745 AA.

AC Y52180;

DT 09-FEB-2000 (first entry)
 XX Human zona pellucida protein B (ZPB) amino acid sequence.
 DE
 XX
 XX zona pellucida protein; ZPA; ZPB; ZPC; oocyte; ovarian cell; antibody;
 KM fusion protein; vaccine; contraceptive; fertilization; fertility;
 KW immunoc contraceptive.
 XX
 XX Homo sapiens.
 OS
 XX
 XX US5989550-A.
 PN
 XX
 XX 23-NOV-1999.
 PD
 XX
 XX 07-JUN-1995; 95US-0480150.
 PF
 XX
 XX 09-NOV-1993; 93US-0149223.
 PR 09-NOV-1992; 92US-0973341.
 PR 29-JAN-1993; 93US-0012990.
 XX
 XX (ZONA-) ZONAGEN INC.
 PA
 XX
 XX Podolski JS, Hsu KT, Harris JD;
 PI
 XX
 DR MPI: 2000-022762/02.
 DR N-PSDB; Z37821.
 XX
 PT Isolated and purified mammalian zona pellucida proteins useful in
 PT methods of immunoc contraceptive -
 XX
 XX Example 10; Column 133-136; 84pp; English.
 XX
 CC This is the human zona pellucida protein B (ZPB) amino acid sequence.
 CC The invention relates to isolated and purified zona pellucida proteins
 CC from mammals, specifically canine, feline and bovine ZPA or their
 CC biologically active fragments also porcine and cynomolgus monkey ZPA. The
 CC zona pellucida is a complex matrix surrounding the mammalian oocyte,
 CC formed of glycoproteins secreted by ovarian cells. The previously
 CC mentioned ZPA proteins can be used to produce a fusion protein to
 CC stimulate production of antibodies in a mammal that recognize a ZPA
 CC polypeptide of the mammal. The ZPA polypeptides are useful as vaccines
 CC for selectively effecting transient infertility in mammals. ZP has an
 CC important role in maintaining the oocyte and in oocyte-sperm interactions
 CC and so is a target for a contraceptive agent which interferes with the
 CC fertilization process. Providing a specific polypeptide reduces the need
 CC for a complex mixture of many ZP proteins which may create unwanted and
 CC harmful side effects. The duration of transient fertility is controllable
 CC and can be maintained in a controllable and/or predictable fashion.
 CC
 CC
 SO Sequence 745 AA;

Query Match 54.4%; Score 43; DB 21; Length 745;
 Best Local Similarity 70.0%; Pred. No. 47;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 GRLCVQDGFEV 13
 I I I I I I I
 Db 369 gelctgdgfm 378

RESULT 14
 ID Y52984 standard; Protein; 745 AA.
 XX
 AC Y52984;
 XX
 XX 14-FEB-2000 (first entry)
 DT
 XX Human zona pellucida ZPA protein sequence.
 DE
 XX zona pellucida; ZPA; ZPB; ZPC; infertility; sterility; vaccine;
 KM immunoc contraceptive; contraceptive; immunisation.

XX
 OS Homo sapiens.
 XX
 XX US5976545-A.
 PN
 XX
 XX 02-NOV-1999.
 PD
 XX
 XX 07-JUN-1995; 95US-0484158.
 PF
 XX
 XX 09-NOV-1993; 93US-0149223.
 PR 09-NOV-1992; 92US-0973341.
 PR 29-JAN-1993; 93US-0012990.
 XX
 XX (ZONA-) ZONAGEN INC.
 PA
 XX
 XX Podolski JS, Hsu KT, Harris JD;
 PI
 XX
 DR MPI: 2000-037080/03.
 DR N-PSDB; Z33271.
 XX
 PT New recombinant mammalian zona pellucida C proteins, for induction of
 PT female sterility -
 XX
 XX Example 11; Column 133-136; 87pp; English.
 XX
 CC The present invention describes recombinant zona pellucida (ZP) proteins,
 CC specifically ZPC proteins. Also described are fusion proteins of ZPC with
 CC any of keyhole limpet haemocyanin, muramyl dipeptide, histidine tag,
 CC beta-galactosidase or palmitic acid, capable of stimulating mammalian
 CC antibodies that recognise mammalian ZPC. Recombinant ZPC proteins are
 CC used to immunise animals to induce transient infertility or permanent
 CC female sterility in humans or other mammals. Z33243 to Z33254, Z33270 to
 CC Z33274 and Z33285 encode mammalian ZP proteins given in Y52571 to Y52988
 CC from the present invention. Z33255 to Z33269 and Z33275 to Z33284
 CC represent oligonucleotides used in the exemplification of the present
 CC invention.
 CC
 CC
 SO Sequence 745 AA;

Query Match 54.4%; Score 43; DB 21; Length 745;
 Best Local Similarity 70.0%; Pred. No. 47;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 GRLCVQDGFEV 13
 I I I I I I I
 Db 369 gelctgdgfm 378

RESULT 15
 ID Y30424 standard; Protein; 84 AA.
 XX
 AC Y30424;
 XX
 XX 15-NOV-1999 (first entry)
 DT
 XX Mature nematode extracted anticoagulant protein Acanthamoeba
 DE
 XX Nematode extracted anticoagulant protein; NAP; anticoagulant;
 KM serine protease inhibitor; NAP domain; factor VIIa/TF.
 XX
 XX Ancylostoma caninum.
 OS
 XX
 XX US5955294-A.
 PN
 XX
 XX 21-SEP-1999.
 PD
 XX
 XX 19-APR-1996; 96US-0634641.
 PF
 XX
 XX 19-APR-1996; 96US-0634641.
 PR 19-APR-1996; 96US-0634641.
 PR 18-OCT-1994; 94US-0326110.
 PR 05-JUN-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0465380.
 PR 05-JUN-1995; 95US-0486397.
 PR 05-JUN-1995; 95US-0486399.
 PR 17-OCT-1995; 95MO-US13231.

PA (CORV-) CORVAS INT INC.

PI Bergum PM, Gansemans YGJ, Jespers LS, Laroche YR;
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PH;
 PI Vlasuk GP;

DR WPI: 1999-539569/45.

PT Screening an isolated protein for Nematode-extracted Anticoagulant
 PT Protein domains

PS Disclosure; Columns 135-138; 197pp; English.

CC The present sequence represents a nematode extracted anticoagulant
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine
 CC protease inhibitor. The protein contains at least one NAP domain which
 CC has selective inhibitory activity for factor VIIa/TF. The specification
 CC describes a method for screening an isolated protein at least one domain
 CC for factor VIIa/TF selective inhibitory activity. The method comprises
 CC determining the time to clotting effected by a concentration of the
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
 CC activated partial thromboplastin time (APTT) assay; calculating
 CC prolongation of clotting effected by the isolated protein in each of
 CC the PT and APTT assay, with respect to a baseline clotting value for
 CC each assay, where prolongation of clotting is calculated as fold
 CC elevation of clotting time relative to a baseline clotting value, where
 CC a doubling of clotting time is deemed a two-fold elevation; and
 CC calculating a PT to APTT prolongation ratio, where a ratio at least
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is
 CC useful for determining if a protein has factor VIIa/TF inhibitory
 CC activity.

SO Sequence 84 AA;

Query Match 53.8%; Score 42.5; DB 20; Length 84;

Best Local Similarity 75.0%; Pred. No. 6.5;

Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
 QY 2 CIGRLCYOD-GF 12
 |||||
 Db 46 CIGRLCYOD-GF 57

Search completed: June 13, 2001, 14:25:51
 Job time: 670 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:23:13 : Search time 87.97 Seconds

(without alignments)
10.937 Million cell updates/sec

Title: PCT-US01-05825A-33

Perfect score: 79

Sequence: 1 FCIGRLCYVDGFWT 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: PIR.67:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	399	2	B82197
2	79	100.0	399	2	A43864
3	43	54.4	676	2	A45984
4	43	54.4	745	1	A48833
5	42	53.2	256	2	T16805
6	41	51.9	155	2	G72580
7	41	51.9	195	2	C72112
8	41	51.9	309	1	B64041
9	41	51.9	329	2	G81299
10	41	51.9	715	2	S70397
11	41	51.9	716	2	S70398
12	40.5	51.3	1209	2	T46027
13	40	50.6	451	2	T24018
14	40	50.6	475	2	T01352
15	40	50.6	631	2	T05868
16	40	50.6	713	1	A34782
17	40	50.6	1235	1	VG1HMJ
18	40	50.6	1324	1	VG1H5J
19	40	50.6	1376	1	VG1HJ2
20	40	50.6	1376	1	U01534
21	39	49.4	44	2	D82364
22	39	49.4	87	2	T39757
23	39	49.4	322	2	S41400
24	39	49.4	340	2	S69194
25	39	49.4	379	2	T41360
26	39	49.4	405	2	T41360
27	38	48.1	245	2	T34189
28	38	48.1	296	2	T26723
29	38	48.1	467	1	VMT77R

30	38	48.1	493	2	E71008	hypothetical prote
31	38	48.1	497	2	S66834	probable membrane
32	38	48.1	698	1	JX0202	long-chain fatty-a
33	38	48.1	805	2	S64238	kinesin-related pr
34	38	48.1	1157	2	T19187	hypothetical prote
35	38	48.1	2946	2	T15840	hypothetical prote
36	37.5	47.5	331	2	E70367	phosphoribosylform
37	37	46.8	161	2	S61389	small basic protei
38	37	46.8	243	2	T34190	hypothetical prote
39	37	46.8	309	2	B82460	hypothetical prote
40	37	46.8	386	2	A82284	conserved hypothet
41	37	46.8	432	2	G64133	adenylosuccinate s
42	37	46.8	468	2	D70557	probable 1gt prote
43	37	46.8	524	2	JN0606	ATP-stimulated glu
44	37	46.8	613	2	T42671	hypothetical prote
45	37	46.8	713	2	S70434	zona pellucida gly

ALIGNMENTS

RESULT 1

B82197

zona occludens toxin VC1458 [imported] - Vibrio cholerae (group O1 strain N16961)

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence-revision 20-Aug-2000 #text-change 15-Sep-2000

C:Accession: B82197

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R. Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, L. R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: B82197

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-399 <HEI>

A:Cross-references: GB:AE004224; GB:AE003852; NID:99655952; PIDN:AAF94615.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1458

A:Map position: 1

Query Match 100.0%; Score 79; DB 2; Length 399;

Best Local Similarity 100.0%; Pred. No. 5.8e-06;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FCIGRLCYVDGFWT 14

DB 288 FCIGRLCYVDGFWT 301

RESULT 2

A43864

zona occludens toxin - Vibrio cholerae

C:Species: Vibrio cholerae

C>Date: 10-Mar-1993 #sequence-revision 18-Nov-1994 #text-change 08-Oct-1999

C:Accession: A43864

R:Baudry, B.; Fasano, A.; Ketley, J.; Kapur, J.B.

Infect. Immun. 60, 428-434, 1992

A>Title: Cloning of a gene (zot) encoding a new toxin produced by Vibrio cholerae.

A:Reference number: A43864; MUID:92112300

A:Accession: A43864

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-399 <BAU>

A:Cross-references: GB:M83563; NID:9155314; PIDN:AAA27582.1; PID:9155315

A>Note: sequence extracted from NCBI backbone (NCBIN:77488, NCBIP:77491)

Query Match 100.0%; Score 79; DB 2; Length 399;

Best Local Similarity 100.0%; Pred. No. 5.8e-06;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FCIGRLCVDGFEVT 14
| | | | | | | | | | | | | | | | | |
Db 288 FCIGRLCVDGFEVT 301

RESULT 3

A45984

sperm-binding glycoprotein ZP2 precursor - rabbit (fragment)

N:Alternate names: 75k zona pellucida glycoprotein; zona pellucida protein 2; zona pellucida

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 03-May-1994 #sequence_revision 05-Jan-1996 #text_change 05-Jan-1996

C/Accession: A45984

R:Lee, V.H.; Schnoebel, E.; Prasad, S.; Cheung, P.; Timmons, T.M.; Cook, R.; Dunbar, B.S.

J. Biol. Chem. 268, 12412-12417, 1993

A:Title: Identification and structural characterization of the 75-kDa rabbit zona pellucida

A:Reference number: A45984; MUID:93286072

A:Accession: A45984

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-676 <LEE>

A:Cross-references: GB:L12167

C:Superfamily: sperm-binding glycoprotein ZP2; ZP domain homology

C:Keywords: glycoprotein; oocyte; transmembrane protein

F:329-369/Domain: ZP domain homology <ZPH>

Query Match 54.4%; Score 43; DB 2; Length 676;

Best Local Similarity 70.0%; Pred. No. 13;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 GRLCVDGFEV 13
| | | | | | | | | | | | | | | | | |
Db 327 GELCTODGFM 336

RESULT 4

A48833

sperm-binding glycoprotein ZP2 precursor - human

N:Alternate names: 75k zona pellucida glycoprotein; zona pellucida protein 2; zona pellucida

C:Species: Homo sapiens (man)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Sep-1999

C/Accession: A48833

R:Liang, L.F.; Dean, J.

Dev. Biol. 156, 399-408, 1993

A:Title: Conservation of mammalian secondary sperm receptor genes enables the promoter C

A:Reference number: A48833; MUID:93215931

A:Accession: A48833

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-745 <LTA>

A:Cross-references: GB:M90366; NID:g292939; PIDN:AAA61335.1; PID:g292940

A:Experimental source: ovary

A:Note: sequence extracted from NCBI backbone (NCBIN:129165, NCBI:P.129166)

C:Genetics:

A:Gene: GDB:2P2

A:Cross-references: GDB:6278878; OMIM:182888

C:Superfamily: sperm-binding glycoprotein ZP2; ZP domain homology

C:Keywords: glycoprotein; oocyte; transmembrane protein

F:371-631/Domain: ZP domain homology <ZPH>

Query Match 54.4%; Score 43; DB 1; Length 745;

Best Local Similarity 70.0%; Pred. No. 15;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 GRLCVDGFEV 13
| | | | | | | | | | | | | | | | | |
Db 369 GELCTODGFM 378

RESULT 5

T16805
hypothetical protein T05A7.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C/Accession: T16805

R:Chissoe, S.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid T05A7.

A:Reference number: T16805

A:Accession: T16805

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-256 <CHI>

A:Cross-references: EMBL:U40028; NID:g1055143; PID:g1055150; PIDN:AAA81120.1; CESP:T0

C:Genetics:

A:Gene: CESP:T05A7.7

A:Insertions: 18/2; 43/2; 119/1; 192/1; 220/1

Query Match 53.2%; Score 42; DB 2; Length 256;

Best Local Similarity 50.0%; Pred. No. 7.9;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 FCIGRLCVDGFEVT 14
| | | | | | | | | | | | | | | | | |
Db 204 YCIGGICVQNPMT 217

RESULT 6

G72580

hypothetical protein APE1926 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C/Accession: G72580

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatakeyama, Y.; Jin-no, K.; Ta

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A:Reference number: A72450; MUID:99310339

A:Accession: G72580

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-155 <KAM>

A:Cross-references: DBJ:AP000062; NID:g5105244; PIDN:BAA80932.1; PID:d1044718; PID:g

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1926

Query Match 51.9%; Score 41; DB 2; Length 155;

Best Local Similarity 60.0%; Pred. No. 7.4;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 CIGRLCVDG 11
| | | | | | | | | | | | | | | | | |
Db 118 CIGRLCARTG 127

RESULT 7

C72112

hypothetical protein CP0606 (imported) - Chlamydia pneumoniae (strains CWL029 and

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000

C/Accession: C72112; D81559

R:Kallman, S.; Mitchell, W.; Marathe, R.; Lamell, C.; Fan, J.; Olinger, L.; Grilwood,

Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606

A:Accession: C72112

A:Molecule type: DNA

A:Residues: 1-195 <ARN>

A:Cross-references: GB:AE001603; GB:AE001363; NID:g4376422; PIDN:AAD18318.1; PID:g437

A:Experimental source: strain CWL029

R:Read, T.D.; R:runham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of *Chlamydia trachomatis* Moyn and *Chlamydia pneumoniae* AR39.

A:Reference number: AB1500; MUID:20150255

A:Accession: DB1559

A:Molecule type: DNA

A:Residues: 1-195 <REA>

A:Cross-references: GB:AE002218; GB:AE002161; NID:g7189509; PIDN:AAF38422.1; PID:g7189555

A:Experimental source: strain AR39, HL cells

A:Genetics:

A:Gene: Cpn0165; Cpn0606

C:Accession: T46027
 R:Benes, V.; Wurmback, E.; Drzonek, H.; Ansoerge, W.; Meyes, H.W.; Lemcke, K.; Mayer, K.F.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: 223019
 A:Accession: T46027
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1209 <BEN>
 A:Cross-references: EMBL:AL132977
 A:Experimental source: cultivar Columbia; BAC clone T10K17
 C:Genetics:
 A:Map position: 3
 A:Insertions: 31/3; 154/3; 241/2; 270/2; 317/3; 441/3; 485/3; 528/3
 A:Note: T10K17.260
 C:Superfamily: Arabidopsis thaliana hypothetical protein T10K17.260

Query Match 51.3%; Score 40.5; DB 2; Length 1209;
 Best Local Similarity 64.3%; Pred. No. 61;
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

OY 2 CI-GRICVODGEV 14
 ||: ||| | |||:
 Db 407 CVHGRLVVGDGEVS 420

RESULT 13
 T24018
 hypothetical protein R07B7.11 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T24018
 R:Harris, B.
 submitted to the EMBL Data Library, July 1996
 A:Reference number: Z19830
 A:Accession: T24018
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-451 <HLL>
 A:Cross-references: EMBL:Z75955; P1DN:CAB00120.1; GSPDB:GN00023; CESP:R07B7.11
 A:Experimental source: clone R07B7
 C:Genetics:
 A:Gene: CESP:R07B7.11
 A:Map position: 5
 A:Insertions: 61/1; 107/3; 317/3; 351/3
 C:Superfamily: alpha-galactosidase

Query Match 50.6%; Score 40; DB 2; Length 451;
 Best Local Similarity 45.5%; Pred. No. 30;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 FCIGRLCYODG 11
 ||: || | |:
 Db 9 FCVGAFCIDNG 19

RESULT 14
 T01352
 hypothetical protein F6N15.2 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 14-May-1999
 C:Accession: T01352
 R:Ryan, E.; Edwards, J.; Pape, K.
 submitted to the EMBL Data Library, May 1998
 A:Description: The sequence of A. thaliana F6N15.
 A:Reference number: Z14297
 A:Accession: T01352
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-475 <RYA>
 A:Cross-references: EMBL:AF069229; NID:g3193311; PID:g3193319
 A:Experimental source: cultivar Columbia

C:Genetics:
 A:Map position: 4
 A:Insertions: 194/1
 A:Note: F6N15.2

Query Match 50.6%; Score 40; DB 2; Length 475;
 Best Local Similarity 50.0%; Pred. No. 31;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 3 IGRICVODGE 12
 ||: ||| | ||:
 Db 206 VSRICVEDGEY 215

RESULT 15
 T05868
 hypothetical protein T29A15.120 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 24-Nov-1999
 C:Accession: T05868
 R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.;
 submitted to the Protein Sequence Database, March 1999
 A:Reference number: Z15455
 A:Accession: T05868
 A:Molecule type: DNA
 A:Residues: 1-631 <BEV>
 A:Cross-references: EMBL:AL035602
 A:Experimental source: cultivar Columbia; BAC clone T29A15
 C:Genetics:
 A:Map position: 4
 A:Insertions: 76/3; 108/3; 146/2; 211/3; 285/2; 319/1; 354/1; 394/2; 437/3; 464/1; 493/3
 A:Note: T29A15.120
 C:Superfamily: Arabidopsis thaliana hypothetical protein T29A15.120

Query Match 50.6%; Score 40; DB 2; Length 631;
 Best Local Similarity 54.5%; Pred. No. 41;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 FCIGRLCYODG 11
 ||: ||| | ||:
 Db 131 FCVGEICKNG 141

Search completed: June 13, 2001, 14:23:14
 Job time: 745 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:30:38 ; Search time 51.57 Seconds
(without alignments)
9.300 Million cell updates/sec

Title: PCT-US01-05825A-33
Perfect score: 79
Sequence: 1 FCIGRLCYDDGFVT 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	399	1	P38442 vibrio chol
2	43	54.4	666	1	P48829 oryctolagus
3	43	54.4	745	1	005996 homo sapien
4	43	54.4	745	1	077726 macaca radi
5	41	51.9	309	1	P44298 haemophilus
6	41	51.9	715	1	P47983 canis famli
7	41	51.9	716	1	P47984 felis silve
8	40	50.6	713	1	P20239 mus musculu
9	40	50.6	1120	1	068428 rat cytomeg
10	40	50.6	1235	1	P11225 murine coro
11	40	50.6	1324	1	P11224 murine coro
12	40	50.6	1376	1	P22432 murine coro
13	40	50.6	1376	1	002385 murine coro
14	39	49.4	87	1	007764 schizosacch
15	39	49.4	340	1	047898 flavobacter
16	39	49.4	511	1	Q9YMS8 diatrophila
17	38	48.1	158	1	Q09571 caenorhabdi
18	38	48.1	467	1	P02898 trypanosoma
19	38	48.1	664	1	P54440 rhizobium s
20	38	48.1	698	1	P31321 homo sapien
21	38	48.1	699	1	P41212 homo sapien
22	38	48.1	805	1	P33086 saccharomyc
23	38	48.1	1157	1	P50286 caenorhabdi
24	37	46.8	431	1	P90747 caenorhabdi
25	37	46.8	524	1	064516 mus musculu
26	37	46.8	524	1	063060 rattus norv
27	37	46.8	716	1	P42099 sus scrofa
28	37	46.8	725	1	P52709 caenorhabdi
29	37	46.8	775	1	015041 homo sapien
30	37	46.8	775	1	SM3E_MOUSE
31	37	46.8	876	1	BGL1_SACFI
32	37	46.8	1012	1	DPOL_HSV7J
33	37	46.8	1171	1	DPOL_HSV7I

34	37	46.8	1171	1	DPOL_HSV72	O9Yus2 herpesvirus
35	37	46.8	1224	1	RBP2_YEAST	P08518 saccharomyc
36	37	46.8	2210	1	RPO_TACYC	P20430 tacaribe vi
37	36	45.6	310	1	V084_HSV7J	P52534 human herpe
38	36	45.6	482	1	GABD_ECOLI	P25526 escherichia
39	36	45.6	497	1	N04M_ACACA	Q37325 acanthamoeb
40	36	45.6	643	1	SYRM_YEAST	P38714 saccharomyc
41	36	45.6	796	1	YRS7_CAEEL	Q10003 caenorhabdi
42	36	45.6	880	1	BGL2_SACFI	P22507 saccharomyc
43	36	45.6	1212	1	UTY_MOUSE	P79457 mus musculu
44	36	45.6	2470	1	TOR1_YEAST	P35169 saccharomyc
45	35	44.3	132	1	RR8_PINTH	P41634 pinus thunb

ALIGNMENTS

RESULT	ID	ZOT_VIBCH	STANDARD:	PRT:	399 AA.
AC	P38442	09L706	Q9R3V6		
DT	01-OCT-1994	(Rel. 30, Created)			
DT	01-OCT-2000	(Rel. 40, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	ZONA OCCUDENS TOXIN (ZONULAR OCCUDENS TOXIN).				
GN	ZOT OR VC1458.				
OS	Vibrio cholerae.				
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.				
OX	NCBI_TaxID=666;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CLASSICAL, INABA 569B;				
RX	MEDLINE=92112300; PubMed=1730472;				
RA	Baudry B., Fasano A., Kelsey J., Kaper J.B.;				
RT	"Cloning of a gene (zot) encoding a new toxin produced by Vibrio cholerae.";				
RL	Infect. Immun. 60:428-434(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=KNH002;				
RA	Shin H.J., Park Y.C., Kim Y.C.;				
RT	"Cloning and nucleotide sequence analysis of the virulence gene cassette from Vibrio cholerae KNH002 isolated in Korea.";				
RL	Misailumunrag Hoji 35:205-210(1999).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=O139-TOR OGANA;				
RA	Zhi-Yong H., Wei-Jie Z., Xiang-Fu W.;				
RT	"Cloning and Expression of zot Gene from Vibrio cholerae.";				
RL	Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=EL TOR 86015 / SEROTYPE O1;				
RA	Kan B., Liu Y.Q., Qi G.M., Gao S.Y.;				
RT	Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=EL TOR N16961 / SEROTYPE O1;				
RX	MEDLINE=20406833; PubMed=10952301;				
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,				
RA	Dodson R.R., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,				
RA	Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,				
RA	Ermlaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,				
RA	McDonald L., Uteback T., Fleishmann R.D., Nierman W.C., White O.,				
RA	Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,				
RA	Fraser C.M.;				
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.";				
RL	Nature 406:477-483(2000).				
RN	[6]				
RP	CHARACTERIZATION.				
RX	MEDLINE=91271365; PubMed=2052603;				
RA	Fasano A., Baudry B., Pumpilin D.W., Wasserman S.S., Tall B.D.,				

```

RA RA Ketley J.M., Kaper J.B.: "Vibrio cholerae produces a second enterotoxin, which affects
RT intestinal tight junctions."
RL Proc. Natl. Acad. Sci. U.S.A. 88:5242-5246(1991).
CC CC -1- FUNCTION: INCREASES THE PERMEABILITY OF THE SMALL INTESTINE MUCOSA
CC BY AFFECTING THE STRUCTURE OF INTERCELLULAR TIGHT JUNCTIONS
CC (ZONULA OCCLUDENS).
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
-----
DR DR EMBL, M83563; AAA27582.1; -
DR DR EMBL, AF175708; AAD51358.1; -
DR DR EMBL, AF123049; AAD26854.1; -
DR DR EMBL, AF220606; AAF29547.1; -
DR DR EMBL, AE004224; AAF94615.1; -
DR DR PIR: A43864; A43864.
DR DR TIGR: VC1458; -.
KW Enterotoxin; Toxin.
FT FT VARIANT 45 45 M -> I (IN STRAIN 569B) .
FT FT VARIANT 100 100 V -> A (IN STRAINS 569B AND 86015) .
FT FT VARIANT 272 272 V -> A (IN STRAIN 569B) .
FT FT VARIANT 281 281 V -> A (IN STRAIN 569B) .
FT FT VARIANT 349 349 V -> S (IN STRAIN 86015) .
FT FT VARIANT 381 381 K -> R (IN STRAIN 86015) .
FT FT CONFLICT 386 399 INTENDKGLGNSIF -> VKKEKESTIKSFL (IN REF.
SQ SEQUENCE 399 AA; 44903 MW; 3C7424B758176774 CRC64;
Query Match 100.0%; Score 79; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 2,7e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FCIGRLCVODGEFVT 14
Db 288 FCIGRLCVODGEFVT 301
|||||
|||
RESULT 2
ZP2_RABIT STANDARD; PRT; 666 AA.
ID ZP2_RABIT
AC P48829;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 2 (ZONA PELLUCIDA GLYCOPROTEIN
DE ZP2) (ZONA PELLUCIDA PROTEIN A) (75 KDA ZONA PELLUCIDA PROTEIN)
DE (FRAGMENT) .
GN ZP2 OR ZPA.
OS Oryctolagus cuniculus (Rabbit) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND WHITE; TISSUE=Ovary;
RX MEDLINE=93286072; PubMed=7685342;
RA Lee V.H., Schwoebel E.D., Prasad S.V., Cheung P., Timmons T.M.,
RA Cook R.G., Dunbar B.S.;
RT "Identification and structural characterization of the 75-kDa rabbit
RT zona pellucida protein."
RL J. Biol. Chem. 268:12412-12417(1993).
CC CC -1- FUNCTION: ZP2 FORMS WITH ZP1 AND ZP3 THE ZONA PELLUCIDA, IN
CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC ZP2 ACTS AS A SECONDARY SPERM RECEPTOR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
CC MATRIX.
```

```
CC -1- PM: PROTEOLYTICALLY CLEAVED AFTER FERTILIZATION, AND THIS  
CC MODIFICATION, ALONG WITH PRESUMED CHANGES IN ZP3 MAY PLAY AN  
CC IMPORTANT ROLE IN THE POSTFERTILIZATION BLOCK TO POLYSPERMY.  
CC -1- PM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES  
CC (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
DR EMBL; L12167; AAA31502.1; -.  
DR InterPro; IPR001507; -.  
DR Pfam; PF00100; zona_pellucida; 1.  
DR ProSite; PS00682; ZP_DOMAIN; 1.  
KW Glycoprotein; Sulfation; Sperm; Receptor; Transmembrane;  
KW Extracellular matrix.  
FT CHAIN 1 666 ZONA PELLUCIDA SPERM-BINDING PROTEIN 2,  
FT DOMAIN <1 636 EXTRACELLULAR (POTENTIAL).  
FT TRANSMM 637 656 POTENTIAL..  
FT DOMAIN 657 666 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 318 585 ZP.  
FT CARBOHD 38 38 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHD 126 126 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHD 171 171 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHD 613 613 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 666 AA; 73644 MM; DCC8B2BA2D21020B CRC64;
```

```
Oy 4 GRLCYODGEV 13      54.4%; Score 43; DB 1; Length 666;  
| | | | | Pred. No. 6;  
Db 317 GELCTQDSFM 326    Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
RESULT 3  
ZP2_HUMAN STANDARD; PRT; 745 AA.  
ID ZP2_HUMAN AC 005396;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 2 PRECURSOR (ZONA PELLUCIDA  
DE GLYCOPROTEIN ZP2) (ZONA PELLUCIDA PROTEIN A).  
GN ZP2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Arthropoda; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RX MEDLINE=93215931; PubMed=8385033;  
RA Liang L.-F., Dean J.;  
RT "Conservation of mammalian secondary sperm receptor genes enables the  
RT promoter of the human gene to function in mouse oocytes.";  
RN Dev. Biol. 156:399-408(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99425270; PubMed=10438293;  
RA Lotfus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,  
RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
```

RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
 RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.:
 RT "Genome duplications and other features in 12 Mb of DNA sequence from
 human chromosome 16p and 16q.";
 RL Genomics 60:295-308(1999).
 CC -1- FUNCTION: ZP2 FORMS WITH ZP1 AND ZP3 THE ZONA PELLUCIDA, IN
 CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
 CC ZP2 ACTS AS A SECONDARY SPERM RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXTRACELLULAR
 CC MATRIX.
 CC -1- TISSUE SPECIFICITY: OOCYTES.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING THE 2-WEEK GROWTH PHASE OF
 CC OOGENESIS, PRIOR TO OVULATION.
 CC -1- PPM: IS PROTEOLYTICALLY CLEAVED AFTER FERTILIZATION, AND THIS
 CC MODIFICATION, ALONG WITH PRESUMED CHANGES IN ZP3 MAY PLAY AN
 CC IMPORTANT ROLE IN THE POSTFERTILIZATION BLOCK TO POLYSPERM.
 CC -1- PPM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M90366; AAA61335.1; -
 DR EMBL: AF001550; AAB67599.1; -
 DR PIR: A48833; A48833.
 DR MIM: 182888; -
 DR InterPro: IPR001507; -
 DR Pfam: PF00100; zona_pellucida; 1.
 DR PRINTS: PR00023; ZPELUCIDA.
 DR PROSITE: PS00682; ZP_DOMAIN; 1.
 KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;
 KW Extracellular matrix;
 FT SIGNAL 1 38
 FT CHAIN 39 745 ZONA PELLUCIDA SPERM-BINDING PROTEIN 2.
 FT DOMAIN 39 716 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 717 736 POTENTIAL.
 FT DOMAIN 737 745 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 370 637 ZP.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 745 AA; 82356 MW; 2614DA79738F2CDD CRC64;
 Query Match 54.4%; Score 43; DB 1; Length 745;
 Best Local Similarity 70.0%; Pred. No. 6.7;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 4 GRLCVDGDFV 13
 DB 369 GELCTODGFM 378
 RESULT 4
 ID ZP2_MACRA STANDARD; PRT; 745 AA.
 AC 077726;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 2 PRECURSOR (ZONA PELLUCIDA
 DE GLYCOPROTEIN ZP2) (ZONA PELLUCIDA PROTEIN A).
 GN ZP2.
 OS Macaca radiata (Bonnet monkey).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 CC Cercopithecoidea; Macaca.
 CC NCBI_TaxID=9548;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=98250422; PubMed=9590540;
 RA Jethanandani P., Santhanam R., Gupta S.K.;
 RT "Molecular cloning and expression in Escherichia coli of cDNA
 RT encoding bonnet monkey (Macaca radiata) zona pellucida
 RT glycoprotein-zp2.";
 RL Mol. Reprod. Dev. 50:229-239(1998).
 CC -1- FUNCTION: ZP2 FORMS WITH ZP1 AND ZP3 THE ZONA PELLUCIDA, IN
 CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
 CC ZP2 ACTS AS A SECONDARY SPERM RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXTRACELLULAR
 CC MATRIX.
 CC -1- PPM: IS PROTEOLYTICALLY CLEAVED AFTER FERTILIZATION, AND THIS
 CC MODIFICATION, ALONG WITH PRESUMED CHANGES IN ZP3 MAY PLAY AN
 CC IMPORTANT ROLE IN THE POSTFERTILIZATION BLOCK TO POLYSPERM.
 CC -1- PPM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y10690; CAA71693.1; -
 DR InterPro: IPR001507; -
 DR Pfam: PF00100; zona_pellucida; 1.
 DR PRINTS: PR00023; ZPELUCIDA.
 DR PROSITE: PS00682; ZP_DOMAIN; 1.
 KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;
 KW Extracellular matrix;
 FT SIGNAL 1 38
 FT CHAIN 39 745 ZONA PELLUCIDA SPERM-BINDING PROTEIN 2.
 FT DOMAIN 39 716 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 717 736 POTENTIAL.
 FT DOMAIN 737 745 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 370 637 ZP.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 745 AA; 82710 MW; 731D9AFA4D3EE028 CRC64;
 Query Match 54.4%; Score 43; DB 1; Length 745;
 Best Local Similarity 70.0%; Pred. No. 6.7;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 4 GRLCVDGDFV 13
 DB 369 GELCTODGFM 378
 RESULT 5
 ID YBCK_HAEIN STANDARD; PRT; 309 AA.
 AC P44298;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOTHETICAL PROTEIN H11730.

GN H11730.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.
 RX NCBI_Taxid=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RD / KW20 / ATCC 51907;
 RA MEDLINE=95350630; PubMed=7542800;
 RA Fitchman R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 McKenney K., Sutton G., Fitzhugh W., Fields C.A., Goeyne J.D.,
 Scott J.D., Shirley R., Liu L.-T., Glodex A., Kelley J.M.,
 Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 Ueberback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
 Fine L.D., Fritchman J.L., Geoghegan N.S.M.,
 Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 Venter J.C.,
 RA *Whole-genome random sequencing and assembly of Haemophilus
 RT Influenzae Rd.;
 RT Science 269:496-512(1995).
 CC -1- SIMILARITY: STRONG, TO E.COLI YBKG.
 CC -1- SIMILARITY: TO B.SUBTILIS YCSJ AND YEAST UREA AMIDOLYASE (DURL.2).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32845; AAC23376.1; -.
 DR TIGR: H11730; -.
 DR Hypothetical protein.
 KM SEQUENCE 309 AA; 34562 MW; E34D87B4C838EFID CRC64;
 SO

Query Match 51.9%; Score 41; DB 1; Length 309;
 Best Local Similarity 80.0%; Pred. No. 6.3;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4. GRLCYODGFV 13
 I I I I I I I I
 DB 111 GYLCOVGDFV 120

RESULT 6
 ZP2_CANFA STANDARD; PRT; 715 AA.
 AC P47983;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 2 PRECURSOR (ZONA PELLUCIDA
 DE GLYCOPROTEIN ZP2) (ZONA PELLUCIDA PROTEIN A).
 GN ZP2 OR ZPA.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Carnivora; Fissipedia; Canidae; Canis.
 OC NCBI_Taxid=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=95143578; PubMed=7841460;
 RA Harris J.D., Hibler D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,
 Sacco A.G.;
 RA "Cloning and characterization of zona pellucida genes and cDNAs from
 RT a variety of mammalian species: the ZPA, ZPB and ZPC gene families.";
 RL DNA Seq. 4:361-393(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RA Okazaki Y., Isojima S., Sugimoto M.;

RA Okazaki Y., Isojima S., Sugimoto M.;
 RL Submitted (Jan-1995) to the EMBL/Genbank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U05779; AA074386.1; -.
 DR EMBL: D45069; BA08097.1; -.
 DR Interpro: IPR001507; -.
 DR Pfam: PF00100; zona_pellucida; 1.
 DR PRINTS: PR00023; zpellucida.
 DR PROSITE: PS00682; ZP.DOMAIN; 1.
 KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;
 KM Extracellular matrix.
 FT SIGNAL 1
 FT CHAIN 39 715
 FT DOMAIN 39 684
 FT TRANSMEM 685 705
 FT DOMAIN 706 715
 FT DOMAIN 367 634
 FT CARBOHYD 87 87
 FT CARBOHYD 193 193
 FT CARBOHYD 220 220
 FT CARBOHYD 266 266
 FT CARBOHYD 321 321
 FT CONFLICT 15 15
 FT CONFLICT 292 292
 FT CONFLICT 328 328
 FT CONFLICT 599 599
 FT SEQUENCE 715 AA; 79938 MW; 508D6DE804F4DC5C CRC64;
 SO

Query Match 51.9%; Score 41; DB 1; Length 715;
 Best Local Similarity 70.0%; Pred. No. 14;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4. GRLCYODGFV 13
 I I I I I I I I
 DB 366 GDLCTODGFM 375

RESULT 7
 ZP2_FEICA STANDARD; PRT; 716 AA.
 AC P47984;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 2 PRECURSOR (ZONA PELLUCIDA
 DE GLYCOPROTEIN ZP2) (ZONA PELLUCIDA PROTEIN A).
 GN ZP2 OR ZPA.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Carnivora; Fissipedia; Felidae; Felis.
 OC NCBI_Taxid=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=95143578; PubMed=7841460;
 RA Harris J.D., Hibler D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,
 Sacco A.G.;
 RA "Cloning and characterization of zona pellucida genes and cDNAs from
 RT a variety of mammalian species: the ZPA, ZPB and ZPC gene families.";
 RL DNA Seq. 4:361-393(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RA Okazaki Y., Isojima S., Sugimoto M.;

RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ZP2 FORMS WITH ZP1 AND ZP3 THE ZONA PELLUCIDA, IN
CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC ZP2 ACTS AS A SECONDARY SPERM RECEPTOR.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
CC MATRIX.
CC -1- PM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES
CC (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U05776; AAA74388.1; -
DR EMBL: D45067; BAA08095.1; -
DR InterPro: IPR001507; -
DR Pfam: PF00100; zona_pellucida; 1.
DR PRINTS: PR00023; ZPPELLUCIDA.
DR PROSITE: PS00682; ZP_DOMAIN: 1.
KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;
KW Extracellular matrix.
FT SIGNAL 1 38 BY SIMILARITY.
FT CHAIN 39 716 ZONA PELLUCIDA SPERM-BINDING PROTEIN 2.
FT DOMAIN 39 686 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 687 707 POTENTIAL.
FT DOMAIN 708 716 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 369 636 ZP.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 397 397 V -> G (IN REF. 2).
FT CONFLICT 483 483 L -> P (IN REF. 2).
FT CONFLICT 637 637 F -> S (IN REF. 2).
FT SEQUENCE 716 AA; 80135 MW; C574549682CB671 CRC64;
SQ
Query Match 51.98; Score 41; DB 1; Length 716;
Best Local Similarity 70.08; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 4 GRLCVDGDFV 13
DB 368 GDLCTODGFM 377
RESULT 8
ZP2_MOUSE
ID ZP2_MOUSE STANDARD; PRT; 713 AA.
AC P20239;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 2 PRECURSOR (ZONA PELLUCIDA
DE GLYCOPROTEIN ZP2) (ZONA PELLUCIDA PROTEIN A).
GN ZP2 OR ZP-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 35-50 AND 428-437.
RC TISSUE-Ovary;
RA MEDLINE=90205829; PubMed=1690843;
RX Liang L.-F., Chamow S.M., Dean J.;
RT "Oocyte-specific expression of mouse Zp-2: developmental regulation
RT of the zona pellucida genes.";

RL Mol. Biol. 10:1507-1515(1990).
CC -1- FUNCTION: ZP2 FORMS WITH ZP1 AND ZP3 THE ZONA PELLUCIDA, IN
CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC ZP2 ACTS AS A SECONDARY SPERM RECEPTOR.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
CC MATRIX.
CC -1- TISSUE SPECIFICITY: OOCYTES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING THE 2-WEEK GROWTH PHASE OF
CC OOGENESIS, PRIOR TO OVULATION.
CC -1- PM: PROTEOLYTICALLY CLEAVED AFTER FERTILIZATION. AND THIS
CC MODIFICATION, ALONG WITH PRESUMED CHANGES IN ZP3 MAY PLAY AN
CC IMPORTANT ROLE IN THE POSTFERTILIZATION BLOCK TO POLYSPERMY.
CC -1- PM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES
CC (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M34148; AAA40586.1; -
DR PIR: A34782; A34782.
DR MGD: MGI:99214; Zp2.
DR InterPro: IPR001507; -
DR Pfam: PF00100; zona_pellucida; 1.
DR PRINTS: PR00023; ZPPELLUCIDA.
DR PROSITE: PS00682; ZP_DOMAIN: 1.
KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;
KW Extracellular matrix.
FT SIGNAL 1 34
FT CHAIN 35 713 ZONA PELLUCIDA SPERM-BINDING PROTEIN 2.
FT DOMAIN 35 683 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 684 703 POTENTIAL.
FT DOMAIN 704 713 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 363 630 ZP.
FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 264 264 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 713 AA; 80209 MW; DCF9AE6CCD3461BF CRC64;
SQ
Query Match 50.68; Score 40; DB 1; Length 713;
Best Local Similarity 63.68; Pred. No. 21;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 3 IGRLCVODGDFV 13
DB 361 IDELCADGDFM 371.
RESULT 9
DPOIL_RCMVM
ID DPOIL_RCMVM STANDARD; PRT; 1120 AA.
AC 085428;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA POLYMERASE (EC 2.7.7.7).
GN ULS4.
OS Rat cytomegalovirus (strain Mastricht).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Murinegaleovirus.
OC NCBI_TaxID=79700;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96335691; PubMed=8757999;

RA Beuken E., Slobbe R., Bruggeman C.A., Vink C.;
 RT "Cloning and sequence analysis of the genes encoding DNA polymerase,
 RT glycoprotein B, ICPI8.5 and major DNA-binding protein of rat
 RT cytomegalovirus." 77:1559-1562(1996).
 CC J. Gen. Virol. 77:1559-1562(1996).
 CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
 CC N PYROPHOSPHATE + DNA(N).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, U50550; AAC56433.1; -.
 CC InterPro: IPR002064; -.
 CC Pfam: PF00136; DNA_POL.B: 1.
 CC PRINTS: PR00106; DNAPOLB.
 CC PROSITE: PS00116; DNA_POLYMERASE_B: 1.
 CC DR transferase: DNA-directed DNA polymerase; DNA replication;
 CC KW DNA-binding; Nuclear protein.
 CC SQ SEQUENCE 1120 AA; 124853 MW; 12FFAE95EF54FC4 CRC64;
 OY 3 IGRICVQDG 11
 OY :|:|||||
 DB 492 VGKICVQDG 500
 Query Match 50.6%; Score 40; DB 1; Length 1120;
 Best Local Similarity 66.7%; Pred. No. 32;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 10
 VGL2_CVMJH STANDARD; PRT: 1235 AA.
 ID VGL2_CVMJH
 AC P11225;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN)
 DE [CONTAINS: SPIKE PROTEIN S1 (90B); SPIKE PROTEIN S2 (90A)].
 GN S.
 OS Murine coronavirus MHV (strain JHM).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OC NCBI_TaxID=11144;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=87111467; PubMed=3072248;
 RX Schmidt I., Skinner M.A., Siddell S.G.;
 RT "Nucleotide sequence of the gene encoding the surface projection
 RT glycoprotein of coronavirus MHV-JHM.";
 RT J. Gen. Virol. 68:47-56(1987).
 RL -1- FUNCTION: THE PEPLOMER PROTEIN MEDIATES THE BINDING OF VIRIONS
 CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION
 CC AND IN SYNCTIUM FORMATION.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, X04797; CAA28484.1; -.
 CC DR EMBL, D00093; BAA00057.1; -.

DR PIR, A33095; VGJHM.
 DR InterPro: IPR002552; -.
 DR Pfam: PF01601; Corona_S2; 2.
 KW Glycoprotein; Envelope protein; Transmembrane; Signal.
 FT SIGNAL 1 10
 FT CHAIN 1 1235
 FT CHAIN 11 628
 FT CHAIN 629 1235
 FT CHAIN 11 1174
 FT DOMAIN 1175 1197
 FT TRANSFM 1198 1235
 FT DOMAIN 1198 1215
 FT DOMAIN 31 31
 FT CARBOHYD 31 31
 FT CARBOHYD 60 60
 FT CARBOHYD 134 134
 FT CARBOHYD 192 192
 FT CARBOHYD 357 357
 FT CARBOHYD 435 435
 FT CARBOHYD 536 536
 FT CARBOHYD 568 568
 FT CARBOHYD 576 576
 FT CARBOHYD 599 599
 FT CARBOHYD 648 648
 FT CARBOHYD 665 665
 FT CARBOHYD 804 804
 FT CARBOHYD 1091 1091
 FT CARBOHYD 1101 1101
 FT CARBOHYD 1120 1120
 FT CARBOHYD 1136 1136
 FT CARBOHYD 1157 1157
 SQ SEQUENCE 1235 AA; 136653 MW; 25962ADB6C1F92DD2 CRC64;
 OY 2 CIGRLCVQDG 11
 OY :||:|:|
 DB 762 CIGSTCAEDG 771
 Query Match 50.6%; Score 40; DB 1; Length 1235;
 Best Local Similarity 60.0%; Pred. No. 35;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 11
 VGL2_CVMAS STANDARD; PRT: 1324 AA.
 ID VGL2_CVMAS
 AC P11224;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN)
 DE [CONTAINS: SPIKE PROTEIN S1 (90B); SPIKE PROTEIN S2 (90A)].
 GN S.
 OS Murine coronavirus MHV (strain A59).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OC NCBI_TaxID=11142;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=88072088; PubMed=2825419;
 RX Luytjes W., Sturman L.S., Bredendijk P.J., Charite J.,
 RA van der Zelfst B.A.M., Horzinek M.C., Spaan W.J.M.;
 RT "Primary structure of the glycoprotein E2 of coronavirus MHV-A59 and
 RT identification of the trypsin cleavage site.";
 RT Virology 161:479-487(1987).
 RL -1- FUNCTION: THE PEPLOMER PROTEIN MEDIATES THE BINDING OF VIRIONS
 CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION
 CC AND IN SYNCTIUM FORMATION.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M18379; AAA46455.1; -

DR PIR; A27402; VG1H59.

DR InterPro: IPR002552; -

DR Pfam: PF01601; Corona_S2; 2.

KM Glycoprotein; Envelope protein; Transmembrane; Signal.

FT SIGNAL 1 16

FT CHAIN 17 1324 E2 GLYCOPROTEIN.

FT CHAIN 17 717 SPIKE PROTEIN S1.

FT CHAIN 718 1324 SPIKE PROTEIN S2.

FT DOMAIN 17 1265 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1266 1286 POTENTIAL.

FT DOMAIN 1287 1324 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 1287 1304 CYS-RICH.

FT CARBOHYD 31 31 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 657 657 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 665 665 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 688 688 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 737 737 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 734 734 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 893 893 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1180 1180 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1190 1190 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1209 1209 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1225 1225 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1246 1246 N-LINKED (GLCNAC. . .) (POTENTIAL).

SO SEQUENCE 1324 AA; 146019 MW; E198EF8F0BCDBF0E CRC64;

Query Match 50.6% Score 40; DB 1; Length 1324;
Best Local Similarity 60.0% Pred. NO. 38;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CIGRLCVODG 11
111 1 : 11
Db 851 CIGSTCAEDG 860

RESULT 12
VGL2_CVM4 STANDARD; PRT; 1376 AA.

AC P22432;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN) [CONTAINS: SPIKE PROTEIN S1 (90B); SPIKE PROTEIN S2 (90A)].

GN S

OS Murine coronavirus MHV (strain wild type 4) (MHV-4).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;

OC Coronaviridae; Coronavirus.

OX NCBI_TaxID:12760;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-9005815; PubMed-2556846;
RA Parker S.E., Gallagher T.M., Buchmeier M.J.;
RT "Sequence analysis reveals extensive polymorphism and evidence of deletions within the E2 glycoprotein gene of several strains of murine hepatitis virus."
RL Virology 173:664-673(1989).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE-9135381; PubMed-1966429;
RA Parker S.E., Buchmeier M.J.;

RT "RNA sequence analysis of the E2 genes of wildtype and neuroattenuated mutants of MHV-4 reveals a hypervariable domain.";
RL Adv. Exp. Med. Biol. 276:395-402(1990).

CC -1- FUNCTION: THE PELOMER PROTEIN MEDIATES THE BINDING OF VIRIONS TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION AND IN SYNCTYL FORMATION.

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN

CC -1- SIMILARITY: NEARLY IDENTICAL TO THE E2 GLYCOPROTEINS FROM MHV-JHM AND MHV-A59 STRAINS, EXCEPT FOR AN N-TERMINAL INSERTION.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; M32789; AAA46456.1; -

CC EMBL; S51114; AB19590.1; -

DR PIR; A33748; VG1HJ2.

DR InterPro: IPR002552; -

DR Pfam: PF01601; Corona_S2; 2.

KM Glycoprotein; Envelope protein; Transmembrane; Signal.

FT SIGNAL 1 14

FT CHAIN 15 1376 E2 GLYCOPROTEIN.

FT CHAIN 15 769 SPIKE PROTEIN S1.

FT DOMAIN 15 1376 SPIKE PROTEIN S2.

FT DOMAIN 1321 1320 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1339 1338 POTENTIAL.

FT DOMAIN 1339 1376 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 429 599 CYS-RICH.

FT CARBOHYD 31 31 IMPORTANT FOR THE NEUROVIRULENCE.

FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 582 582 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 709 709 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 717 717 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 740 740 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 789 789 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 806 806 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 945 945 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1232 1232 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1242 1242 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1261 1261 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1277 1277 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1298 1298 N-LINKED (GLCNAC. . .) (POTENTIAL).

SO SEQUENCE 1376 AA; 151882 MW; 8BC01B978252094E CRC64;

Query Match 50.6% Score 40; DB 1; Length 1376;
Best Local Similarity 60.0% Pred. NO. 39;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CIGRLCVODG 11
111 1 : 11
Db 903 CIGSTCAEDG 912

RESULT 13
VGL2_CVM4C STANDARD; PRT; 1376 AA.

AC Q02385;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN) [CONTAINS: SPIKE PROTEIN S1 (90B); SPIKE PROTEIN S2 (90A)].

GN S.
 OS Murine coronavirus MHV (strain JHMV / variant CL-2).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=33735;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9226864; PubMed=116938;
 RA Taguchi F., Ikeda T., Shida H.;
 RT "Molecular cloning and expression of a spike protein of neurotropic
 RT murine coronavirus JHMV variant CL-2.";
 RL J. Gen. Virol. 73:1065-1072(1992).
 CC -1- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
 CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION
 CC AND IN SYNCTIUM FORMATION.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, D10235; BAA01085.1; -.
 DR PIR: J01534; J01534.
 DR InterPro: IPR002552; -.
 DR Pfam: PF01601; Corona_S2; 2.
 KW Glycoprotein; Envelope protein; Transmembrane; Signal.
 FT SIGNAL 1 14
 FT CHAIN 15 1376 E2 GLYCOPROTEIN.
 FT CHAIN 15 769 SPIKE PROTEIN S1.
 FT CHAIN 770 1376 SPIKE PROTEIN S2.
 FT DOMAIN 15 1330 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1321 1338 POTENTIAL.
 FT DOMAIN 1339 1356 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 31 31 CYS-RICH.
 FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 709 709 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 717 717 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 789 789 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 806 806 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 945 945 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1232 1232 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1242 1242 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1261 1261 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1277 1277 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1298 1298 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1376 1376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1376 AA; 152041 MW; 98C30DD979F9E75 CRC64;

Query Match 50.6%; Score 40; DB 1; Length 1376;
 Best Local Similarity 60.0%; Pred. No. 39;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 CIGRLCVDG 11
 DB 903 CIGSTCAEDG 912

RESULT 14
 RS21_SCHPO STANDARD; PRT; 87 AA.
 AC P05764; 094496;
 DT 01-NOV-1988 (rel. 09; Created)
 DT 01-OCT-2000 (rel. 40; Last sequence update)

DT 01-OCT-2000 (rel. 40; Last annotation update)
 DE 40S RIBOSOMAL PROTEIN S21 (S28).
 GN RPS21 OR SPBCL8E5.06.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=86104253; PubMed=3910104;
 RA Itoh T., Otake E., Matsui K.A.;
 RT "Primary structures of ribosomal protein YS25 from Saccharomyces
 RT cerevisiae and its counterparts from Schizosaccharomyces pombe and
 RT rat liver";
 RL Biochemistry 24:7418-7423(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Wood V., Barrall B.G., Rajandream M.A., Xiang Z., Aves S.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE S21E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AL035077; CAA22666.1; -.
 DR PIR: B23862; B23862.
 DR InterPro: IPR001931; -.
 DR Pfam: PF01249; Ribosomal_S21e; 1.
 DR PROSITE: PS00996; RIBOSOMAL_S21E; 1.
 KW Ribosomal protein; Acetylation.
 FT MOD_RES 1 1 ACETYLATION.
 FT CONFLICT 37 37 C -> A (IN REF. 1).
 FT CONFLICT 41 41 A -> Q (IN REF. 1).
 FT CONFLICT 68 68 C -> D (IN REF. 1).
 FT CONFLICT 73 73 T -> Q (IN REF. 1).
 FT CONFLICT 80 80 E -> T (IN REF. 1).
 SQ SEQUENCE 87 AA; 9617 MW; 970F7B10DC348B6A CRC64;

Query Match 49.4%; Score 39; DB 1; Length 87;
 Best Local Similarity 58.3%; Pred. No. 4.1;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 CIGRLCVDGDFV 13
 DB 68 CINRLTDDGLL 79

RESULT 15
 ASPG_FLAME STANDARD; PRT; 340 AA.
 AC 047898;
 DT 01-NOV-1997 (rel. 35; Last sequence update)
 DT 01-NOV-1997 (rel. 35; Last sequence update)
 DT 30-MAY-2000 (rel. 39; Last annotation update)
 DE N4 (BETA-N-ACETYLGLUCOSAMINYL)-L-ASPARAGINASE PRECURSOR (EC 3.5.1.26)
 DE (GLUCOSYLASPARAGINASE) (ASPARTYLGLUCOSAMINIDASE) (N4-(N-ACETYL-BETA-
 DE GLUCOSAMINYL)-L-ASPARAGINE AMIDASE) (AGA).
 OS Flavobacterium meningosepticum.
 OC Bacteria; CFB group; Flavobacteriia; Flavobacteriaceae;
 OC Chryseobacterium.
 OX NCBI_TaxID=238;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC STRAIN=ELDER / ATCC 33958;
 RX MEDLINE=95142655; PubMed=7840643;

RA Tarentino A.L., Quinones G., Hauer C.R., Changchien L.-M.,
RA Plummer T.H. Jr.;
RT "Molecular cloning and sequence analysis of *Flavobacterium*
RT *meningosepticum* glycosylasparaginase: a single gene encodes the alpha
RT and beta subunits.";
RL Arch. Biochem. Biophys. 316:399-406(1995).
RN (2)
RP SEQUENCE OF 46-59 AND 197-211.
RX MEDLINE-94071939; PubMed-8250923;
RA Tarentino A.L., Plummer T.H. Jr.;
RT "The first demonstration of a procaryotic glycosylasparaginase.";
RL Blochem. Biophys. Res. Commun. 197:179-186(1993).
RN (3)
RX X-RAY CRYSTALLOGRAPHY (2.32 ANGSTROMS).
RP MEDLINE-98200483; PubMed-9541410;
RA Xuan J., Tarentino A.L., Grimwood B.G., Plummer T.H. Jr., Cui T.,
RA Guan C., van Rooy P.;
RT "Crystal structure of glycosylasparaginase from *Flavobacterium*
RT *meningosepticum*.";
RL Protein Sci. 7:774-781(1998).
CC -1- FUNCTION: CLEAVES THE GLYC-ASN BOND WHICH JOINS OLIGOSACCHARIDES
CC TO THE PEPTIDE OF ASPARAGINE-LINKED GLYCOPROTEINS. REQUIRES THAT
CC THE GLYCOSYLATED ASPARAGINE MOIETY IS NOT SUBSTITUTED ON
CC ITS AMINO (R1) AND CARBOXYL (R2) TERMINUS.
CC -1- CATALYTIC ACTIVITY: N4-(BETA-N-ACETYL-D-GLUCOSAMINYL)-L-ASPARAGINE
CC + H(2)O -> N-ACETYL-BETA-GLUCOSAMINYLAMINE + L-ASPARTATE.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAINS.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE ASPARAGINASE 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U08028; AAA68868.1; -;
DR PDB: 1AVY; 29-APR-98.
DR MEROPS: T02.001; -;
DR InterPro: IPR000246; -;
DR Pfam: PF01112; Asparaginase-2; 1.
KW Signal: Hydrolase; Periplasmic; 3d-structure.
FT SIGNAL 1 45
FT CHAIN 46 196 GLYCOSYLASPARAGINASE, ALPHA CHAIN.
FT CHAIN 197 340 GLYCOSYLASPARAGINASE, BETA CHAIN.
FT DISULFID 113 126
FT DISULFID 213 277
FT ACT_SITE 197 197 OR 315.
SQ SEQUENCE 340 AA; 37262 MW; 4C36E5061B4E53D7 CRC64;

Query Match 49.4%; Score 39; DB 1; Length 340;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 GRLCVODGF 12
1 :|||||
DB 312 GAYCICODGF 320

Search completed: June 13, 2001, 14:30:39
Job time: 529 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: June 13, 2001, 14:29:47 ; Search time 150.43 Seconds
(without alignments)
10.908 Million cell updates/sec

Title: PCT-US01-05825A-33
Perfect score: 79
Sequence: 1 FCIGRLCVDGDFVT 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPREMBL_15:
1: SP-archaea:
2: SP-bacteria:
3: SP-fungi:
4: SP-human:
5: SP-invertebrate:
6: SP-mammal:
7: SP-mhc:
8: SP-organelle:
9: SP-phage:
10: SP-plant:
11: SP-rodent:
12: SP-unclassified:
13: SP-vertebrate:
14: SP-virus:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	323	2	Q918F5 vibrio mim
2	79	100.0	399	2	Q9R3V6 vibrio chol
3	79	100.0	399	2	Q917G6 vibrio chol
4	43	54.4	127	10	Q9S1C5 arabidopsis
5	43	54.4	436	2	Q9R3G6 arabidopsis
6	42	53.2	256	5	Q22208 caenorhabd
7	42	53.2	263	10	Q9ZQ11 arabidopsis
8	41	51.9	155	1	Q9YAL6 aeropyrum p
9	41	51.9	195	2	Q9Z919 chlamydia p
10	41	51.9	329	2	Q9PMEO campylobact
11	41	51.9	549	11	Q9WU64 mus musculu
12	41	51.9	611	2	Q6S514 mycobacteri
13	41	51.9	870	3	Q14424 coccidioid
14	40.5	1209	10	Q9M2P3 arabidopsis	
15	40	50.6	138	5	Q9NLK9 leishmania
16	40	50.6	315	2	Q9WU16 synechococ
17	40	50.6	451	5	Q21801 caenorhabd
18	40	50.6	475	10	Q81325 arabidopsis
19	40	50.6	631	10	Q9T087 arabidopsis

Result No.	Score	Query Match	Length	DB ID	Description
20	40	50.6	1223	14	Q9WR70 murine hepa
21	40	50.6	1324	14	Q39227 murine hepa
22	40	50.6	1324	14	Q913E7 murine hepa
23	40	50.6	1376	14	Q83330 murine hepa
24	40	50.6	1546	4	Q9NS27 homo sapien
25	40	50.6	1551	4	Q75445 homo sapien
26	39	49.4	44	2	Q9KVN1 vibrio chol
27	39	49.4	87	3	Q94496 schizosacch
28	39	49.4	340	5	Q9YMS8 drosophila
29	39	49.4	379	2	P72738 synechocyst
30	39	49.4	405	3	P87241 schizosacch
31	39	49.4	1007	14	Q91F17 porcine cyt
32	39	49.4	3781	5	Q9V1S5 drosophila
33	38	48.1	83	5	Q76927 drosophila
34	38	48.1	128	10	Q9S118 arabidopsis
35	38	48.1	217	4	Q14964 homo sapien
36	38	48.1	245	5	Q18925 caenorhabd
37	38	48.1	296	5	Q9XX28 caenorhabd
38	38	48.1	470	5	P90594 trypanosoma
39	38	48.1	493	1	Q59088 pyrococcus
40	38	48.1	497	3	Q08280 saccharomyc
41	38	48.1	513	2	Q914N2 streptococ
42	38	48.1	787	5	Q9UA44 aplysia cal
43	38	48.1	1007	14	Q91F16 porcine cyt
44	38	48.1	1007	14	Q91F15 porcine cyt
45	38	48.1	1007	14	Q91F13 porcine cyt

ALIGNMENTS

RESULT 1
Q918F5 PRELIMINARY; PRT; 323 AA.

AC Q918F5; 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DI 01-OCT-2000 (TREMBLrel. 15, last annotation update)
DE ZOT (FRAGMENT).

GN ZOT
OS Vibrio mimicus.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=674;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=PT5;
RX MEDLINE=20143766; PubMed=10678967;
RA Boyd E.F., Moyer K.E., Shi L., Waldor M.K.;
RT "Infectious CTXphi and the vibrio pathogenicity island prophage in
RT vibrio mimicus: evidence for recent horizontal transfer between V.
RL mimicus and V. cholerae.";
DR Infect. Immun. 68:1507-1513(2000).
EMBL AF207857; AAF40142.1; --

FT NON_TER 1
FT SEQUENCE 323 AA; 36306 MW; 01C12DAE9B873C3B CRC64;
SQ

Query Match 100.0%; Score 79; DB 2; Length 323;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FCIGRLCVDGDFVT 14
Db 237 FCIGRLCVDGDFVT 250

RESULT 2
Q9R3V6 PRELIMINARY; PRT; 399 AA.
AC Q9R3V6; 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)

DT 01-OCT-2000 (Tremblrel. 15, last annotation update)
 DE ZONULAR OCCUDENS TOXIN (ZONA OCCUDENS TOXIN).
 GN ZOT OR VC1458.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SRRAIN-KN1H002:
 RA Shih H.T., Park Y.C., Kim Y.C.:
 RT "Cloning and nucleotide sequence analysis of the virulence gene
 RL cassette from Vibrio cholerae KN1H002 isolated in Korea.";
 RL Misamunhag Hoji 35:205-210(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SRRAIN-0139-TOR OGMA:
 RA Zhi-Tong H., Wei-Jie Z., Xiang-Fu W.:
 RT "Cloning and Expression of zot Gene from Vibrio cholerae.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SRRAIN-EL TOR N16961 / SRRORTYPE 01:
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heideberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischman R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.:
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 DR EMBL: AF175708; AAD51358.1; -
 DR EMBL: AF123049; AAD26854.1; -
 DR EMBL: AE004224; AAF94615.1; -
 DR TIGR: VC1458; -
 SQ SEQUENCE 399 AA; 44903 MW; 3C7424B758176774 CRC64;

 Query Match 100.0%; Score 79; DB 2; Length 399;
 Best Local Similarity 100.0%; Pred. No. 5.9e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 OY 1 FCIGRLCVDGFEVT 14
 DB 288 FCIGRLCVDGFEVT 301

 RESULT 3
 ID 09L706 PRELIMINARY; PRT; 399 AA.
 AC 09L706;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, last annotation update)
 DE ZOT.
 GN ZOT.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SRRAIN-86015:
 RA Kan B., Liu Y.Q., Qi G.M., Gao S.Y.:
 RT "Vibrio cholerae nct-cxphl whole genome, include rstr(Rstr),
 RL rsta(Rsta), rstr(Rstr), cep(Cep), orfU(OrfU), ace(Ace) and zot(Zot)
 RL genes.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF220606; AAF29547.1; -
 SQ SEQUENCE 399 AA; 44990 MW; CF6A3DBCC9E23EB1 CRC64;

Query Match 100.0%; Score 79; DB 2; Length 399;
 Best Local Similarity 100.0%; Pred. No. 5.9e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 OY 1 FCIGRLCVDGFEVT 14
 DB 288 FCIGRLCVDGFEVT 301

 RESULT 4
 ID 09S1C5 PRELIMINARY; PRT; 127 AA.
 AC 09S1C5;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, last annotation update)
 DE ATG31440 PROTEIN.
 GN ATG31440.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Magnoliophyta; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Eukaryota; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Bentto M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Beil C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Freus D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.:
 RT "Sequence and analysis of chromosome II of Arabidopsis thaliana.";
 RL Nature 402:761-768(1999).
 DR EMBL: AC007169; AAD26475.1; -
 SQ SEQUENCE 127 AA; 14351 MW; F2B8B19352550222 CRC64;

 Query Match 54.4%; Score 43; DB 10; Length 127;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 OY 1 FCIGRLC 7
 DB 51 FCIGRLC 57

 RESULT 5
 ID 09KYG6 PRELIMINARY; PRT; 436 AA.
 AC 09KYG6;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, last annotation update)
 DE PUTATIVE TRANSFERASE.
 GN SCC61A.27.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Oliver K., Harris D.:
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SRRAIN-A3(2);
 RA Bentley S.D., Parkhill J., Barrett B.G., Rajandream M.A.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);
 RA Redenbach M., Kieseer H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinoshita H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL: AL356595; CAB92272.1; -;
 KW Transference.
 SO SEQUENCE 436 AA; 4551 MW; 7A54166F6383854 CRC64;

Query Match 54.4%; Score 43; DB 2; Length 436;
 Best Local Similarity 70.0%; Pred. No. 12;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 CIGRLCVODG 11
 :|||||:
 Db 218 CVGRLCRCKG 227

RESULT 6
 ID Q22208 PRELIMINARY; PRT: 256 AA.

AC Q22208;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
 DE SIMILARITY TO GLYCOPROTEIN H PRECURSOR.
 GN T05A7.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;

RA SEQUENCE FROM N.A.
 RX MEDLINE=94150718; Pubmed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Haller L., Jler M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Lalster N., Latrelle P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Riffen L., Roopra A., Saunders D., Showkeen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sprat J., Wooldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chisoe S.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Waterston R.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U40028; AAA81120.1; -;
 SO SEQUENCE 256 AA; 26045 MW; F55ABB67CF87EB61 CRC64;

Query Match 53.2%; Score 42; DB 5; Length 256;
 Best Local Similarity 50.0%; Pred. No. 10;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 FCIGRLCVODGFV 14
 :|||||:
 Db 204 YCIGICQVNMKST 217

RESULT 7
 ID Q9ZQL1 PRELIMINARY; PRT: 263 AA.

AC Q9ZQL1;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
 DE T13E11.8 PROTEIN.
 GN T13E11.8.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Magnoliopsida; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Eudicotyledons; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA.
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., Vanaken S.E.,
 RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.,
 RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
 RA Fraser C.M., Venter J.C.;
 RT "Arabidopsis thaliana chromosome II BAC T13E11 genomic sequence.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC006217; AAD15523.1; -;
 SO SEQUENCE 263 AA; 29788 MW; F3DB9F4E7BA75B6D CRC64;

Query Match 53.2%; Score 42; DB 10; Length 263;
 Best Local Similarity 53.8%; Pred. No. 11;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 FCIGRLCVODGFV 13
 :|||||:
 Db 57 FIDGRLCMEDPPT 69

RESULT 8
 ID Q9YAL6 PRELIMINARY; PRT: 155 AA.

AC Q9YAL6;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE HYPOTHEETICAL 15.8 KDA PROTEIN APE1926.
 GN APE1926.

OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococaceae;
 OC Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KI;
 RX MEDLINE=99310339; Pubmed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S., Akai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kushida N., Oguchi A., Aoki K., Kubota K., Nakamura Y.,
 RA Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix KI.";
 RL DNA Res. 6:83-101(1999).
 DR EMBL: AF000062; BAA80932.1; -;
 KW Hypothetical protein.
 SO SEQUENCE 155 AA; 15806 MW; CC105579BDC7EEB CRC64;

Query Match 51.9%; Score 41; DB 1; Length 155;
 Best Local Similarity 60.0%; Pred. No. 9.7;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 CIGRLCVODG 11
 :|||||:
 Db 118 CIGRLCARTG 127

```

RESULT 9
ID Q92919 PRELIMINARY; PRT; 195 AA.
AC Q92919;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 23.0 KDA PROTEIN.
GN CPN0165 OR CPJ0165 OR CP0606.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kilmann S., Mitchell W., Marathe R., Lamme C., Fan J., Hyman R.W.,
RT Olinger L., Grimwood J., Davys R.W., Stephens R.S.;
RL "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RL from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Unayen L.A., Uetebach T.,
RA Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,
RA Bowman C., Dodson R., Gwin N., Nelson W., Deboy R., Kolonay J.,
RA McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RL pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL: AEO02545; AAD18318.1; -
DR EMBL: AEO02545; BAA98375.1; -
DR EMBL: AEO02218; AAF38422.1; -
DR TIGR: CP0606; -
DR TIGR: CPJ0165; -
DR TIGR: CPN0165; -
KW Hypothetical protein.
SQ SEQUENCE 195 AA; 23037 MW; D35C244BE5FA4F1 CRC64;

Query Match 51.9%; Score 41; DB 2; Length 195;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 FCIGRLCYODGEVY 14
   11:1 1111:1
DB 74 FCILTRSYODGYLT 87

RESULT 10
O9PME0 PRELIMINARY; PRT; 329 AA.
AC O9PME0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PHOSPHORIBOSYLFORMYLGLYCINAMIDINE CYCLO-LIGASE (EC 6.3.3.1).
GN PURM.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kellly J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
RA Jags K., Karsleyshew A., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
DR EMBL: AL139078; CAB73945.1; -
DR INTERPRO: IPR000728; -
DR INTERPRO: IPR002086; -
DR PFAM: PF00586; ATRS; 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
SQ SEQUENCE 329 AA; 35643 MW; B637ADC82CE5B7F2 CRC64;

Query Match 51.9%; Score 41; DB 2; Length 329;
Best Local Similarity 46.2%; Pred. No. 20;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 FCIGRLCYODGEVY 13
   11:1 11:1
DB 315 FTIGELICNEGIV 327

RESULT 11
O9WU64 PRELIMINARY; PRT; 549 AA.
AC O9WU64;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GLYCEROL KINASE-LIKE PROTEIN 1.
GN GK-RS1 OR GKRS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-129SV; TISSUE=TESTES;
RA Pan Y., Decker W.K., Hug A.H.H.M., Craig W.J.;
RT "Retrospective position of glycerol kinase-related genes from the X
RT chromosome to autosomes. Functional and evolutionary aspects.";
RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF117733; AAD24550.1; -
DR HSSP: P08859; 1GIC.
DR MGD: MGI:891990; GK-rs1.
DR INTERPRO: IPR000577; -
DR INTERPRO: IPR002086; -
DR PFAM: PF00370; EGGY; 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE: PS00445; EGGY_KINASES_2; 1.
DR PROSITE: PS00933; EGGY_KINASES_1; UNKNOWN_1.
KW Kinase.
SQ SEQUENCE 549 AA; 59871 MW; F807404B9F373F9 CRC64;

Query Match 51.9%; Score 41; DB 11; Length 549;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 IGRICVODG 11
   11:1111111
DB 265 VGQLCLVDG 273

RESULT 12
O69514

```

ID 069514 PRELIMINARY; PRT; 611 AA.
 AC 069514;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE DNA POLYMERASE III SUBUNITS GAMMA AND TAU.
 OS DNAX.
 OS Mycobacterium lepre.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OC NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Seeger K., Harris D.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93188700; Pubmed=8446027;
 RA Elgimeier K., Honore N., Woods S.A., Caudron B., Cole S.T.;
 RT "Use of an ordered cosmid library to deduce the genomic organization
 of Mycobacterium lepre."
 RL Mol. Microbiol. 7:197-206(1993).
 DR EMBL: AL023596; CAI:9155.1; -.
 DR INTERPRO: IPR000862; -.
 SO SEQUENCE 611 AA; 65535 MW; 39BAD5F9E636F6 CRC64;

Query Match 51.9%; Score 41; DB 2; Length 611;
 Best Local Similarity 54.5%; Pred. No. 35;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 3 IGRICVODGFEV 13
 Db 186 VGRICAGGV 196

RESULT 13
 ID 014424 PRELIMINARY; PRT; 870 AA.
 AC 014424;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE BETA-GLUCOSIDASE (EC 3.2.1.21).
 GN BGL1.
 OS Coccioides limlitis.
 OC Eukaryota; Fungi; Ascomycota; Onygenales; anamorphic Onygenales;
 OC Coccioides.
 OC NCBI_TaxID=5501;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C735;
 RA Yu J.-J., Thomas P.W., Seshan K., Cole G.T.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U87805; AAB67972.1; -.
 DR INTERPRO: IPR001764; -.
 DR INTERPRO: IPR002772; -.
 DR PRAM: PF00933; Glyco_hydro_3; 1.
 DR PRAM: PF01915; Glyco_hydro_3_C; 1.
 DR PRINTS: PR00133; GLYDRLASE3.
 DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.
 KM Hydrolase; Glycosidase.
 SO SEQUENCE 870 AA; 95467 MW; F0FDDF20CD64E1F CRC64;

Query Match 51.9%; Score 41; DB 3; Length 870;
 Best Local Similarity 72.7%; Pred. No. 49;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FCIGRLCVODG 11
 Db 99 FGIRLCLQDG 109

RESULT 14
 ID 09M2P3 PRELIMINARY; PRT; 1209 AA.
 AC 09M2P3;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE HYPOTHETICAL 136.8 KDA PROTEIN.
 GN T10K17.260.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidops.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Benes V., Wurmback E., Drzonek H., Ansoorge W., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X., Queller F., Salanoubat M.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL132977; CAB67633.1; -.
 KM Hypothetical protein.
 SO SEQUENCE 1209 AA; 136820 MW; 097F949AD98885D CRC64;

Query Match 51.3%; Score 40.5; DB 10; Length 1209;
 Best Local Similarity 64.3%; Pred. No. 82;
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
 QY 2 CI-GRICVODGFEV 14
 Db 407 CVHGRVVGDFVS 420

RESULT 15
 ID 09NLK9 PRELIMINARY; PRT; 138 AA.
 AC 09NLK9;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE HYPOTHETICAL 14.9 KDA PROTEIN (FRAGMENT).
 GN LM26.433.
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OC NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL160493; CAB98051.1; -.
 KM Hypothetical protein.
 FT NON_TER 138
 SO SEQUENCE 138 AA; 14866 MW; 1CFAB56D7E3F08B1 CRC64;

Query Match 50.6%; Score 40; DB 5; Length 138;
 Best Local Similarity 42.9%; Pred. No. 13;
 Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 FCIGRLCVODGFEV 14
 Db 55 FCVARMSYMDGYNT 68

Wed Jun 13 15:01:05 2001

pct-us01-05825a-33.rspt

Search completed: June 13, 2001, 14:29:48
Job time: 549 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:27:09 ; Search time 78.71 Seconds
(without alignments)
3.417 Million cell updates/sec

Title: PCT-US01-05825A-33

Perfect score: 79

Sequence: 1 PCTGRLCVDDGFVT 14

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	400	2	US-08-624-601-8
2	43	54.4	745	1	US-08-453-472-5
3	43	54.4	745	1	US-08-038-948-9
4	43	54.4	745	1	US-08-453-952-5
5	43	54.4	745	2	US-08-484-993B-43
6	43	54.4	745	2	US-08-862-903-5
7	43	54.4	745	2	US-08-484-158B-43
8	43	54.4	745	2	US-08-484-596A-43
9	43	54.4	745	2	US-08-480-150A-43
10	43	54.4	745	2	US-08-458-731-43
11	43	54.4	745	3	US-08-149-223A-43
12	42.5	53.8	84	2	US-08-465-380-51
13	42.5	53.8	84	2	US-08-486-397-51
14	42.5	53.8	84	2	US-08-486-399-51
15	42.5	53.8	84	2	US-08-461-965-51
16	42.5	53.8	84	2	US-08-634-641-51
17	42.5	53.8	84	3	US-09-249-471-51
18	42.5	53.8	84	3	US-09-249-472-51
19	42.5	53.8	84	3	US-09-249-451-51
20	42.5	53.8	84	3	US-08-809-455-51
21	42.5	53.8	84	4	US-09-249-461-51
22	42.5	53.8	84	4	US-09-249-448-51
23	42.5	53.8	162	2	US-08-465-380-64
24	42.5	53.8	162	2	US-08-486-397-64
25	42.5	53.8	162	2	US-08-486-399-64
26	42.5	53.8	162	2	US-08-461-965-64
27	42.5	53.8	162	2	US-08-634-641-64

28	42.5	53.8	162	3	US-09-249-471-64	Sequence 64, Appl
29	42.5	53.8	162	3	US-09-249-472-64	Sequence 64, Appl
30	42.5	53.8	162	3	US-09-249-451-64	Sequence 64, Appl
31	42.5	53.8	162	3	US-08-809-455-64	Sequence 64, Appl
32	42.5	53.8	162	3	US-09-249-461-64	Sequence 64, Appl
33	42.5	53.8	162	4	US-09-249-448-64	Sequence 64, Appl
34	41	51.9	715	2	US-08-484-993B-10	Sequence 10, Appl
35	41	51.9	715	2	US-08-484-158B-10	Sequence 10, Appl
36	41	51.9	715	2	US-08-484-596A-10	Sequence 10, Appl
37	41	51.9	715	2	US-08-480-150A-10	Sequence 10, Appl
38	41	51.9	715	3	US-08-458-731-10	Sequence 10, Appl
39	41	51.9	715	3	US-08-149-223A-10	Sequence 10, Appl
40	41	51.9	716	2	US-08-484-993B-14	Sequence 14, Appl
41	41	51.9	716	2	US-08-484-158B-14	Sequence 14, Appl
42	41	51.9	716	2	US-08-484-596A-14	Sequence 14, Appl
43	41	51.9	716	2	US-08-480-150A-14	Sequence 14, Appl
44	41	51.9	716	3	US-08-458-731-14	Sequence 14, Appl
45	41	51.9	716	3	US-08-149-223A-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-08-624-601-8
Sequence 8, Application US/08624601
Patent No. 5882653
GENERAL INFORMATION:
APPLICANT: Kaper Dr., James B.
TITLE OF INVENTION: Vibrio cholerae O1 (CVD111) and non-O1
TITLE OF INVENTION: CVD112 and CVD112RM serogroup vaccine strains, methods
of making same and products thereof
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Spencer & Frank
STREET: 1100 New York Ave. N.W. Suite 300 East
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/624,601
FILING DATE: 08-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schellier Dr., John W.
REGISTRATION NUMBER: 26,031
REFERENCE/DOCKET NUMBER: BAWC20019P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)414-4000
TELEFAX: (202)414-4040
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Vibrio cholerae
STRAIN: El Tor 7946
IMMEDIATE SOURCE:
CLONE: zot
US-08-624-601-8

Query Match 100.0%; Score 79; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 9.7e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FCIGRLCVODGFMV 14
|||||
Db 288 FCIGRLCVODGFMV 301

RESULT 2

US-08-453-472-5
; Sequence 5, Application US/08453472
; Patent No. 5626846
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453.472
; FILING DATE: 30-May-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/038,948
; FILING DATE: 26-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,462
; FILING DATE: 20-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/364,379
; FILING DATE: 12-JUN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4032 US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 745
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; FEATURE:
; NAME/KEY: ZP2
; LOCATION:

IDENTIFICATION METHOD:
; OTHER INFORMATION: human ZP2 protein
; US-08-453-472-5

Query Match 54.4%; Score 43; DB 1; Length 745;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GRICVODGFMV 13
|||||
Db 369 GELCTODGFMV 378

RESULT 3

US-08-038-948-9
; Sequence 9, Application US/08038948
; Patent No. 564187
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE BASED ON
; TITLE OF INVENTION: ALLOIMMUNIZATION WITH ZONA PELLUCIDA POLYPEPTIDES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/038,948
; FILING DATE: 26-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,462
; FILING DATE: 20-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/364,379
; FILING DATE: 12-JUN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: 99152/E-266-88/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 745 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-038-948-9

Query Match 54.4%; Score 43; DB 1; Length 745;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GRICVODGFMV 13
|||||
Db 369 GELCTODGFMV 378

RESULT 4

US-08-453-952-5
; Sequence 5, Application US/08453952
; Patent No. 5672488
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,952
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/038,948
; FILING DATE: 26-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,462
; FILING DATE: 20-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/364,379
; FILING DATE: 12-JUN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4032 USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 745
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLES:
; FEATURE:
; NAME/KEY: 2P2
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: human 2P2 protein
; US-08-453-952-5

Query Match 54.4%; Score 43; DB 1; Length 745;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 4 GRLCVQDGFV 13
| | | | |

DB 369 GELCTQDGFV 378
RESULT 5
US-08-484-993B-43
; Sequence 43, Application US/08484993B
; Patent No. 5837497
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,993B
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TEXEL: 25-3856
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-993B-43

Query Match 54.4%; Score 43; DB 2; Length 745;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 4 GRLCVQDGFV 13
| | | | |
DB 369 GELCTQDGFV 378

RESULT 6
US-08-862-903-5
; Sequence 5, Application US/08862903
; Patent No. 5916768
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN

Wed Jun 13 15:01:03 2001

pct-us01-05825a-33.rat

Page 4

STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/862,903
FILING DATE: 30-MAY-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/038,948
FILING DATE: 26-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/930,462
FILING DATE: 20-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/364,379
FILING DATE: 12-JUN-1989
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4032 USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 745
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: 2P2
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: human 2P2 protein
US-08-862-903-5
Query Match 54.4%; Score 43; DB 2; Length 745;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

DB 369 GELCTODGFM 378
OY 4 GRUCVODGFV 13
US-08-484-158B-43
Sequence 43, Application US/08484158B
Patent No. 5976545
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.

TITLE OF INVENTION: Pharmaceutical Compositions for
TITLE OF INVENTION: Immunoreception
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray's
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,158B
FILING DATE: 07-JUNE-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 09-NOV-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-92
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 32794
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-158B-43
Query Match 54.4%; Score 43; DB 2; Length 745;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

DB 369 GELCTODGFM 378
OY 4 GRUCVODGFV 13
US-08-484-158B-43
Sequence 43, Application US/08484596A
Patent No. 5981228
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunoreception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,596A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 11-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-596A-43

Query Match 54.4% Score 43; DB 2; Length 745;
Best Local Similarity 70.0% Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 GRLCVQDGFV 13
| | | | |
DB 369 GELCTQDGFV 378

RESULT 9
US-08-480-150A-43
Sequence 43, Application US/08480150A
Patent No. 5989550
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
TITLE OF INVENTION: Materials and Methods for Immuncontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,150A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/149,223
FILING DATE: 09-NOV-1993
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341

FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-150A-43

Query Match 54.4% Score 43; DB 2; Length 745;
Best Local Similarity 70.0% Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 GRLCVQDGFV 13
| | | | |
DB 369 GELCTQDGFV 378

RESULT 10
US-08-458-731-43
Sequence 43, Application US/08458731
Patent No. 6001599
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
TITLE OF INVENTION: Materials and Methods for Immuncontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,731
FILING DATE: 09-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

ADDRESS: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,397
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/269
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-397-51

Query Match 53.8%; Score 42.5; DB 2; Length 84;
Best Local Similarity 75.0%; Pred. No. 2.1;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 2 CIGRLCVOD-GF 12
||||:|||||
Db 46 CIGRVCVDEGF 57

RESULT 14
US-08-486-399-51
Sequence 51, Application US/08486399
Patent No. 5865543
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,399
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-399-51

Query Match 53.8%; Score 42.5; DB 2; Length 84;
Best Local Similarity 75.0%; Pred. No. 2.1;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 2 CIGRLCVOD-GF 12
||||:|||||
Db 46 CIGRVCVDEGF 57

RESULT 15
US-08-461-965-51
Sequence 51, Application US/08461965
Patent No. 5872098
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,965
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.

Wed Jun 13 15:01:03 2001

pct-us01-05825a-33.ra1

Page 8

REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 210/243
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 673510
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-461-965-51

Query Match 53.88; Score 42.5; DB 2; Length 84;
Best Local Similarity 75.0%; Pred. No. 2.1;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Caps 1;

OY 2 CIGRICVQD-GF 12
|||:| | |
Db 46 CIGRVCVDEGF 57 #

Search completed: June 13, 2001, 14:27:10
Job time: 631 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:08:19 ; Search time 229.28 Seconds
(without alignments)
1.995 Million cell updates/sec

Title: PCT-US01-05825a-1

Perfect score: 46

Sequence: 1 GRVCVPG 8

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	8	21	Y79105
2	43	93.5	8	21	Y79109
3	39	84.8	8	21	Y79113
4	39	84.8	8	21	Y79117
5	38	82.6	8	21	Y79106
6	38	82.6	470	14	R34476
7	38	82.6	470	14	R34477
8	38	82.6	470	14	R34478
9	37	80.4	1291	16	R75201
10	36	78.3	8	21	Y79107
11	36	78.3	8	21	Y79121

12	35	76.1	8	21	Y79110	Peptide antagonist
13	35	76.1	399	18	R20006	zonula occludens t
14	35	76.1	720	18	W13574	Mouse beta-3 integ
15	35	76.1	787	18	W13573	Mouse beta-3 integ
16	34	73.9	718	11	R05936	Secreted GPIIb su
17	34	73.9	762	21	Y92448	GPIIb variant Leu
18	34	73.5	762	21	Y92449	GPIIb variant Asp
19	34	73.5	762	21	Y92450	GPIIb variant Ser
20	34	73.9	762	21	Y92451	GPIIb variant Arg
21	34	73.9	762	21	Y92452	GPIIb variant Arg
22	34	73.9	762	21	Y92453	GPIIb variant Cys
23	34	73.9	762	21	Y92454	GPIIb variant Pro
24	34	73.9	762	21	Y92455	GPIIb variant Arg
25	34	73.9	762	21	Y92456	GPIIb variant Ser
26	34	73.9	788	15	R51500	Human platelet GPI
27	34	73.9	788	20	Y49553	Human endothelial
28	34	73.9	788	20	Y49557	Human antithrombin
29	34	73.9	788	21	Y92443	Wild type human GP
30	34	73.9	788	21	Y92444	Variant human GPII
31	34	73.9	788	21	Y81461	Human integrin bet
32	33	71.7	8	21	Y79111	Peptide antagonist
33	33	71.7	13	19	Y20209	Human beta-amyloid
34	33	71.7	26	20	Y02927	Fragment of human
35	33	71.7	31	20	W88384	Human zneu1 EGF-11
36	33	71.7	73	20	W88389	Human zneu1 partia
37	33	71.7	97	22	B50353	Human RIN1f1n pr
38	33	71.7	113	21	G17477	Arabidopsis thalia
39	33	71.7	115	21	B41718	Human ORFX ORF1482
40	33	71.7	141	21	G17476	Arabidopsis thalia
41	33	71.7	153	21	B41638	Human ORFX ORF1402
42	33	71.7	158	20	W88388	Human zneu1 partia
43	33	71.7	169	20	W88380	Human zneu1 partia
44	33	71.7	188	21	Y59439	Human delta3 fragm
45	33	71.7	224	20	Y59870	Human normal uteru

ALIGNMENTS

RESULT 1	
ID	Y79105 standard; Peptide: 8 AA.
Y79105	
XX	Y79105:
AC	
DT	05-JUN-2000 (first entry)
XX	
DE	Peptide antagonist of zonulin.
XX	
KW	zonulin; antagonist; zonula occludens toxin receptor;
KW	blood-brain barrier; antiinflammatory; cerebroprotective;
KW	neuroprotective; dermatological; antitumor; antiviral;
KW	antibacterial; cytostatic; anti-HIV; vulnereary; antiallergic;
KW	hypotensive; immunosuppressive; antiparasitic; vasotropic;
KW	gastrointestinal inflammation; therapy.
OS	Synthetic.
XX	
PN	WO200007609-A1.
XX	
PD	17-FEB-2000.
XX	
PF	28-JUL-1999; 99WO-US16683.
XX	
PR	03-AUG-1998; 98US-0127815.
XX	
PA	(UYMA-) UNIV MARYLAND BALTIMORE.
XX	
PI	Fasano A;
XX	
DR	WPI; 2000-205565/18.
XX	
PT	New peptide antagonist of zonulin useful as antiinflammatory agent for

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
XX shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
XX
PS Claim 1; Page 41; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
CC (2), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical synthesis or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.

XX
XX
XX Sequence 8 AA;

Query Match 100.0%; Score 46; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVCVPG 8
DB 1 grvcvpg 8

RESULT 2
Y79109
ID Y79109 standard; Peptide; 8 AA.

AC Y79109;
XX
XX 05-JUN-2000 (first entry)
XX
XX Peptide antagonist of zonulin.

DE Zonulin; antagonist; zonula occludens toxin receptor;
XX blood-brain barrier; antiinflammatory; cerebroprotective;
XX neuroprotective; dermatological; antitumor; antiviral;
XX antibacterial; cytostatic; anti-HIV; vulnerary; antiallergic;
XX hypotensive; immunosuppressive; antiparasitic; vasotropic;
XX gastrointestinal inflammation; therapy.

XX
XX Synthetic.
XX
XX WO200007609-A1.
XX
XX 17-FEB-2000.
XX
XX 28-JUL-1999; 99WO-US16683.
XX
XX 03-AUG-1998; 98US-0127815.
XX
XX (UYMA-) UNIV MARYLAND BALTIMORE.
XX

PI Fasano A;
XX
XX
XX WPI; 2000-205565/18.

PT New peptide antagonist of zonulin useful as antiinflammatory agent for
XX treating cerebral ischemia, stroke, cerebral edema, gastritis,
XX shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
XX
PS Claim 1; Page 42; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
CC (2), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical synthesis or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.

XX
XX
XX Sequence 8 AA;

Query Match 93.5%; Score 43; DB 21; Length 8;
Best Local Similarity 87.5%; Pred. No. 3.2e+05;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVCVPG 8
DB 1 grvcvpg 8

RESULT 3
Y79113
ID Y79113 standard; Peptide; 8 AA.

AC Y79113;
XX
XX 05-JUN-2000 (first entry)
XX
XX Peptide antagonist of zonulin.

DE Zonulin; antagonist; zonula occludens toxin receptor;
XX blood-brain barrier; antiinflammatory; cerebroprotective;
XX neuroprotective; dermatological; antitumor; antiviral;
XX antibacterial; cytostatic; anti-HIV; vulnerary; antiallergic;
XX hypotensive; immunosuppressive; antiparasitic; vasotropic;
XX gastrointestinal inflammation; therapy.

XX
XX Synthetic.
XX
XX WO200007609-A1.
XX
XX 17-FEB-2000.
XX
XX 28-JUL-1999; 99WO-US16683.
XX
XX

XX	03-AUG-1998;	98US-0127815.
PR		
XX		
PA	(UYMA-) UNIV MARYLAND BALTIMORE.	
XX		
PI	Fasano A;	
DR	WPI; 2000-205565/18.	
PT	New peptide antagonist of zonulin useful as antiinflammatory agent for	
PT	treating cerebral ischemia, stroke, cerebral edema, gastritis,	
PT	shigellosis, viral gastroenteritis, meningitis, encephalomyelitis	
XX		
PS	Claim 1; Page 43; 69pp: English.	
XX		
CC	This present sequence is that of a peptide antagonist of zonulin	
CC	(2), one of 25 such peptides (see Y79105-29) of the invention,	
CC	which bind to a zonula occludens toxin (ZOT) receptor, yet do not	
CC	physiologically modulate the opening of mammalian tight junctions	
CC	(TJ). The peptide antagonists are based on a common motif of ZOT	
CC	and human zonulins, which is believed to be critical for receptor	
CC	binding. They can be prepared by chemical synthesis or by use of	
CC	recombinant DNA techniques. The peptide antagonists are used as an	
CC	antiinflammatory agents in the treatment of gastrointestinal	
CC	inflammation, where they bind to the ZOT receptor in the intestine	
CC	and yet does not physiologically modulate the opening of TJ in the	
CC	intestine. Gastrointestinal inflammation conditions give rise to	
CC	increased intestinal permeability and the peptide is useful for	
CC	treating intestinal conditions that cause protein losing enteropathy	
CC	caused by infection, e.g. Clostridium difficile infection,	
CC	enterocolitis, shigellosis, viral gastroenteritis, parasite	
CC	infestation, bacterial overgrowth, whipple's disease, diseases with	
CC	mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,	
CC	collagenous colitis, inflammatory bowel disease, diseases marked by	
CC	lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,	
CC	sarcoctosis lymphoma, mesenteric tuberculosis, and after surgical	
CC	correction of congenital heart disease with Fontan's operation,	
CC	mucosal diseases without ulceration, e.g. Menetrier's disease,	
CC	eosinophilic gastroenteritis, and immune diseases,	
CC	e.g. systemic lupus erythematosus or food allergies, primarily to	
CC	milk.	
XX		
SO	Sequence	8 AA;
XX		
XX	Query Match	84.8%; Score 39; DB 21; Length 8;
XX	Best Local Similarity	87.5%; Pred. NO. 3.2e+05;
XX	Matches 7; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	1 GRVCVQPG 8	
DB	1 grgcvgpg 8	
XX		
RESULT	4	
Y79117	Y79117 standard; Peptide; 8 AA.	
XX	Y79117;	
XX		
DT	05-JUN-2000 (first entry)	
XX		
DE	Peptide antagonist of zonulin.	
XX		
KW	Zonulin; antagonist; zonula occludens toxin receptor;	
KW	blood-brain barrier; antiinflammatory; cerebroprotective;	
KW	neuroprotective; dermatologic; antitumor; antiviral;	
KW	antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;	
KW	hypotensive; immunosuppressive; antiparasitic; vasotropic;	
KW	gastrointestinal inflammation; therapy.	
XX		
OS	Synthetic.	
XX		

PN	WO200007609-A1.
XX	
PD	17-FEB-2000.
XX	
PF	28-JUL-1999; 99WO-US16683.
XX	
PR	03-AUG-1998; 98US-0127815.
XX	
PA	(UYMA-) UNIV MARYLAND BALTIMORE.
XX	
PI	Fasano A;
XX	
DR	WPI; 2000-205565/18.
XX	
PT	New peptide antagonist of zonulin useful as antiinflammatory agent for
PT	treating cerebral ischemia, stroke, cerebral edema, gastritis,
XX	shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
PS	Claim 1; Page 45; 69pp; English.
XX	
CC	This present sequence is that of a peptide antagonist of zonulin
CC	(2), one of 25 such peptides (see Y79105-29) of the invention,
CC	which bind to a zonula occludens toxin (ZOT) receptor, yet do not
CC	physiologically modulate the opening of mammalian tight junctions
CC	(TJ). The peptide antagonists are based on a common motif of ZOT
CC	and human zonulins, which is believed to be critical for receptor
CC	binding. They can be prepared by chemical synthesis or by use of
CC	recombinant DNA techniques. The peptide antagonists are used as an
CC	antiinflammatory agents in the treatment of gastrointestinal
CC	inflammation, where they bind to the ZOT receptor in the intestine
CC	and yet does not physiologically modulate the opening of TJ in the
CC	intestine. Gastrointestinal inflammation conditions give rise to
CC	increased intestinal permeability and the peptide is useful for
CC	treating intestinal conditions that cause protein losing enteropathy
CC	caused by infection, e.g. Clostridium difficile infection,
CC	enterocolitis, shigellosis, viral gastroenteritis, parasite
CC	infection, bacterial overgrowth, whipple's disease, diseases with
CC	mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC	collagenous colitis, inflammatory bowel disease, diseases marked by
CC	lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC	sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC	correction of congenital heart disease with Fontan's operation,
CC	mucosal diseases without ulceration, e.g. Menetrier's disease,
CC	coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC	e.g. systemic lupus erythematosus or food allergies, primarily to
CC	milk.
XX	
SQ	Sequence 8 AA;
	Query Match 84.8%; Score 39; DB 21; Length 8;
	Best Local Similarity 87.5%; Pred. No. 3.2e+05;
	Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	1 GRVQVPG 8
	1
DB	1 ggvqvpg 8
RESULT 5	
Y79106	
ID	Y79106 standard; Peptide: 8 AA.
XX	
AC	Y79106;
XX	
DT	05-JUN-2000 (first entry)
XX	
DE	Peptide antagonist of zonulin.
XX	
KM	Zonulin: antagonist; zonula occludens toxin receptor;
KM	blood-brain barrier; antiinflammatory; cerebroprotective;
KM	neuroprotective; dermatological; antiviral;
KM	antibacterial; cytostatic; anti-HIV; vulnervary; anti allergic;

PS Claim 4; Page 34-36; 44pp; Japanese.
 XX
 CC cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B
 CC (9405 nucleotides long) and 14 shorter clones were isolated by PCR
 CC amplification, including clone JK4-B. Primer/probes derived from the
 CC sequences of these clones can be used in diagnostic assays for HCV.
 CC See Q40425-Q40439.
 XX
 SQ Sequence 470 AA;
 OY 1 GRVCVOPG 8
 DB 73 grcvcpq 80
 Query Match 82.6%; Score 38; DB 14; Length 470;
 Best Local Similarity 75.0%; Pred. NO. 38;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 RESULT 8
 R34478 ID R34478 standard; Protein; 470 AA.
 XX
 AC R34478;
 XX
 DT 30-JUL-1993 (first entry)
 XX
 DE Encoded by Hepatitis C virus clone JK4-C.
 XX
 KM HCV; non-A, non-B hepatitis virus; NANBHV; liver disease;
 KM polymerase chain reaction; diagnostic method.
 XX
 OS Hepatitis C virus.
 XX
 PN JP05068562-A.
 XX
 PD 23-MAR-1993.
 XX
 PF 30-MAY-1991; 91JP-0153736.
 XX
 PR 30-MAY-1991; 91JP-0153736.
 XX
 PA (SANM) SANMA KAGAKU KENKUSHO CO.
 XX
 DR WPI; 1993-130638/16.
 DR N-PSDB; Q40436.
 XX
 PT DNA and cDNA of hepatitis C virus - useful as probes for
 PT diagnosing HCV infection
 XX
 PS Claim 4; Page 36-38; 44pp; Japanese.
 XX
 CC cDNA was prepared from HCV genomic RNA. Full-length clone JK1-B
 CC (9405 nucleotides long) and 14 shorter clones were isolated by PCR
 CC amplification, including clone JK4-C. Primer/probes derived from the
 CC sequences of these clones can be used in diagnostic assays for HCV.
 CC See Q40425-Q40439.
 XX
 SQ Sequence 470 AA;
 OY 1 GRVCVOPG 8
 DB 73 grcvcpq 80
 Query Match 82.6%; Score 38; DB 14; Length 470;
 Best Local Similarity 75.0%; Pred. NO. 38;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 RESULT 9
 R75201 ID R75201 standard; Protein; 1291 AA.
 XX
 AC R75201;
 XX
 DT 10-MAY-1996 (first entry)
 XX
 DE Tyrosine phosphatase MPTP-delta.
 XX
 KM Tyrosine phosphatase MPTP-delta; murine; brain tissue;
 KM glutathione-S-transferase; fusion protein; E. coli; differentiation;
 KM activation; information transmission; nervous system; immune system;
 KM carcinogenesis.
 XX
 OS Mus musculus domesticus.
 XX
 PN JP07236487-A.
 XX
 PD 12-SEP-1995.
 XX
 PF 28-FEB-1994; 94JP-0054726.
 XX
 PR 28-FEB-1994; 94JP-0054726.
 XX
 PA (TOKS-) TOKYOTO SHINKEI KAGAKU SOGO KENKUSHO ZH.
 XX
 DR WPI; 1995-347455/45.
 DR N-PSDB; Q94311.
 XX
 PT DNA encoding tyrosine phosphatase MPTP delta - useful for
 PT elucidation of signal transmission mechanisms.
 XX
 PS Claim 1; Page 5-11; 14pp; Japanese.
 XX
 CC This sequence represents murine tyrosine phosphatase MPTP-delta. The
 CC cDNA sequence encoding this protein was isolated from murine brain
 CC tissue and was cloned, for expression, into the downstream region of a
 CC glutathione-S-transferase sequence and expressed as a fusion protein
 CC in E. coli. MPTP-delta proteins regulate differentiation and
 CC activation of cells. This sequence can be used in the elucidation of
 CC the molecular mechanism for information transmission in cells,
 CC regulation mechanisms for information transmission in cells,
 CC the mechanism of carcinogenesis.
 XX
 SQ Sequence 1291 AA;
 OY 1 GRVCVOP 7
 DB 1064 grvclop 1070
 Query Match 80.4%; Score 37; DB 16; Length 1291;
 Best Local Similarity 85.7%; Pred. NO. 1.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 RESULT 10
 Y79107 ID Y79107 standard; Peptide; 8 AA.
 XX
 AC Y79107;
 XX
 DT 05-JUN-2000 (first entry)
 XX
 DE Peptide antagonist of zonulin.
 XX
 KM Zonulin; antagonist; zonula occludens toxin receptor;
 KM blood-brain barrier; antiinflammatory; cerebroprotective;
 KM neuroprotective; dermatological; antitumor; antiviral;
 KM antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
 KM hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KM gastrointestinal inflammation; therapy.
 XX
 OS Synthetic.

XX WO200007609-A1.
 XX 17-FEB-2000.
 XX
 XX 28-JUL-1999; 99WO-US16683.
 XX
 XX 03-AUG-1998; 98US-0127815.
 XX
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 XX Fasano A;
 XX
 XX WPI; 2000-205565/18.
 XX
 XX New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 XX Claim 1; Page 41; 69pp; English.

This present sequence is that of a peptide antagonist of zonulin (Z), one of 25 such peptides (see Y79105-29) of the invention, which bind to a zonula occludens toxin (ZOT) receptor, yet do not physiologically modulate the opening of mammalian tight junctions (TJ). The peptide antagonists are based on a common motif of ZOT and human zonulins, which is believed to be critical for receptor binding. They can be prepared by chemical synthesis or by use of recombinant DNA techniques. The peptide antagonists are used as an antiinflammatory agents in the treatment of gastrointestinal inflammation, where they bind to the ZOT receptor in the intestine and yet does not physiologically modulate the opening of TJ in the intestine. Gastrointestinal inflammation conditions give rise to increased intestinal permeability and the peptide is useful for treating intestinal conditions that cause protein losing enteropathy caused by infection, e.g. Clostridium difficile infection, enterocolitis, shigellosis, viral gastroenteritis, parasite infestation, bacterial overgrowth, whipple's disease, diseases with mucosal erosion or ulcerations, e.g. gastritis, gastric cancer, collagenous colitis, inflammatory bowel disease, diseases marked by lymphatic obstruction, e.g. congenital intestinal lymphangiectasia, sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical correction of congenital heart disease with Fontan's operation, CC mucosal diseases without ulceration, e.g. Menetrier's disease, CC coeliac disease, eosinophilic gastroenteritis, and immune diseases, CC e.g. systemic lupus erythematosus or food allergies, primarily to milk.

Sequence 8 AA;

Query Match 78.3%; Score 36; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVCVQPG 8
 ||| ||||
 DB 1 grvlvqpg 8

RESULT 11
 ID Y79121
 Y79121 standard; Peptide; 8 AA.
 AC Y79121;

05-JUN-2000 (first entry)

Peptide antagonist of zonulin.

Zonulin; antagonist; zonula occludens toxin receptor;
 blood-brain barrier; antiinflammatory; cerebroprotective;
 neuroprotective; dermatological; antiulcer; antiviral;

KW antibacterial; cytostatic; anti-HIV; vulnereary; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.

XX Synthetic.
 XX
 XX WO200007609-A1.
 XX
 XX 17-FEB-2000.
 XX
 XX 28-JUL-1999; 99WO-US16683.
 XX
 XX 03-AUG-1998; 98US-0127815.
 XX
 XX (UYMA-) UNIV MARYLAND BALTIMORE.

Fasano A;

WPI; 2000-205565/18.

New peptide antagonist of zonulin useful as antiinflammatory agent for treating cerebral ischemia, stroke, cerebral edema, gastritis, shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 Claim 1; Page 46; 69pp; English.

This present sequence is that of a peptide antagonist of zonulin (Z), one of 25 such peptides (see Y79105-29) of the invention, which bind to a zonula occludens toxin (ZOT) receptor, yet do not physiologically modulate the opening of mammalian tight junctions (TJ). The peptide antagonists are based on a common motif of ZOT and human zonulins, which is believed to be critical for receptor binding. They can be prepared by chemical synthesis or by use of recombinant DNA techniques. The peptide antagonists are used as an antiinflammatory agents in the treatment of gastrointestinal inflammation, where they bind to the ZOT receptor in the intestine and yet does not physiologically modulate the opening of TJ in the intestine. Gastrointestinal inflammation conditions give rise to increased intestinal permeability and the peptide is useful for treating intestinal conditions that cause protein losing enteropathy caused by infection, e.g. Clostridium difficile infection, enterocolitis, shigellosis, viral gastroenteritis, parasite infestation, bacterial overgrowth, whipple's disease, diseases with mucosal erosion or ulcerations, e.g. gastritis, gastric cancer, collagenous colitis, inflammatory bowel disease, diseases marked by lymphatic obstruction, e.g. congenital intestinal lymphangiectasia, sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical correction of congenital heart disease with Fontan's operation, CC mucosal diseases without ulceration, e.g. Menetrier's disease, CC coeliac disease, eosinophilic gastroenteritis, and immune diseases, CC e.g. systemic lupus erythematosus or food allergies, primarily to milk.

Sequence 8 AA;

Query Match 78.3%; Score 36; DB 21; Length 8;
 Best Local Similarity 75.0%; Pred. No. 3.2e+05;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVCVQPG 8
 | : |||||
 DB 1 gglcvqpg 8

RESULT 12
 ID Y79110
 Y79110 standard; Peptide; 8 AA.
 AC Y79110;

05-JUN-2000 (first entry)

Peptide antagonist of zonulin.

KM zonulin; antagonists; zonula occludens toxin receptor;
KM blood-brain barrier; antiinflammatory; cerebroprotective;
KM neuroprotective; dermatologicals; antidote; antiviral;
KM antibacterial; cystostatic; anti-HIV; vulnereary; antiallergic
KM hypotensive; immunosuppressive; antiparasitic; vasotropic;
KM gastrointestinal inflammation; therapy.

OS Synthetic.

PN WO200007609-A1.

PD 17-FEB-2000.

PF 28-JUL-1999; 99WO-US16683.

PR 03-AUG-1998; 98US-0127815.

PA (UYMA-) UNIV MARYLAND BALTIMORE.

PI Fasano A;
vxy

DR WPI; 2000-205565/18

PT New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
XX
PS Claim 1; Page 42; 69pp; English.

Claim 1; Page 42; 69pp; English.

This present sequence is that of a peptide antagonist of zonulin (2), one of 25 such peptides (see Yf9105-29) of the invention, which bind to a zonula occludens toxin (ZOT) receptor, yet do not physiologically modulate the opening of mammalian tight junctions (TJ). The peptide antagonists are based on a common motif of ZOT and human zonulins, which is believed to be critical for receptor binding. They can be prepared by chemical synthesis or by use of recombinant DNA techniques. The peptide antagonists are used as an antiinflammatory agents in the treatment of gastrointestinal inflammation, where they bind the ZOT receptor in the intestine and yet does not physiologically modulate the opening of TJ in the intestine. Gastrointestinal inflammation conditions give rise to increased intestinal permeability and the peptide is useful for treating intestinal conditions that cause protein losing enteropathy caused by infection, e.g. Clostridium difficile infection, enterocolitis, shigellosis, viral gastroenteritis, parasite infestation, bacterial overgrowth, whipple's disease, diseases with mucosal erosion or ulcerations, e.g. gastritis, gastric cancer, collagenous colitis, inflammatory bowel disease, diseases marked by lymphatic obstruction, e.g. congenital intestinal lymphangiectasia, sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical correction of congenital heart disease with Fontan's operation, mucosal diseases without ulceration, e.g. Menetrier's disease, coeliac disease, eosinophilic gastroenteritis, and immune diseases, e.g. systemic lupus erythematosus or food allergies, primarily to milk.

SQ Sequence 8 AA;

Query Match	76.1%	Score 35:	DB 21:	Length 8:	-
Best Local Similarity	75.0%	Pred. No.	3.2e+05,		
Matches 6;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;	

Oy	1	GRVCVQPG	8
		:	
Db	1	grlcvgdg	8

RESULT 13
R20006
ID R20006 standard; Protein; 399 AA

XX AC R20006;

DT 31-MAR-1992 (first entry)
 VV

Zonula occludens toxin.

ZOT; cholera; vaccine; enterotoxin; diarrhoea

OS Vibrio cholerae.

PN WO9118979-A.

PD 12-DEC-1991
xy

PF 05-JUN-1991; 91WO-US03812.
 YY

PR 05-JUN-1990; 90US-05333315.
XX

PA (UYMA-) UNIV MARYLAND BALTI.
YY

PI Kaper JB, Baudry-Maurelli B, Fasano A, YX

DR WPI; 1992-007465/01.
DR N-PSDR: 020185

XX New Vibrato cho

PT Fragment encoding toxin, used as vaccines against cholera

XX
PS Disclosure; Fig 18; 83pp; English.

CC The amo acid sequence is that of the zonula occludens toxin (ZOT) .
CC It may be responsible for diarrhoea in some strains of cholera and
CC the ZOT gene or fragments of it are deleted from strains of *Vibrio*
CC cholerae ('V.c.') to be used as vaccines. These V.c. strains have 100%
CC efficacy in protecting humans against subsequent infection with a
CC strain of a similar serotype and avoid undesirable side effects such
CC as diarrhoea, nausea and cramping. Cultures of these strains may be
CC used for prodn. of vaccines against cholera.

50 Sequence 399 AA;

Query Match	76.1%	Score 35:	DB 13:	Length 399:
Best Local Similarity	75.0%	Pred. No. 1:	le+02:	
Matches	6:	Conservative	1:	Indels 0: Gaps 0

```

QY      1  GRVCVQPG  8
          ||: ||| |
Db      291  grlcvgdq  298

```

RESULT	14
W13574	
ID	W13574 standard; Protein; 720 AA

AC W13574;

DT 03-JUN-1997 (first entry)

DE Mouse beta-3 Integrin (truncated).

KW Beta-3 integrin; bone resorption; osteoporosis; osteoclast;

3 XX 3

XX	XX
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

XX
2
3
4
5
6
7

XX 07-AUG-2005 ACWA-F012005

XX

PR 31-AUG-1995; 95US-0003020.
XX
PA (MERI) MERCK & CO INC.
XX
PI Duong LT, Nult EM, Rodan GA:
XX
DR WPI: 1997-179273/16.
XX N-PSDB: T61291.
XX
PT Mouse full length or truncated beta-3 integrin sub-unit(s) and
PT coding sequences - useful for screening for cpds. that inhibit bone
PT resorption
XX
PS Claim 20; Fig 3A-3B: 30pp; English.
XX
XX A novel truncated mouse vitronectin integrin receptor beta-3 subunit
CC (W13574) lacks the transmembrane and cytoplasmic domains of the
CC full-length receptor (see also W13573) and is considered to be a
CC soluble ligand binding integrin that is able to circulate
CC throughout the organism and to suppress or compete with the normal
CC signals mediated by the wild-type receptor. It can be produced in
CC host cells utilizing a novel cDNA clone (T61291). The full-length
CC and truncated beta-3 integrin subunits can be used in assays to
CC identify novel cpds. which inhibit the bone absorption process,
CC esp. in osteoporosis.
SQ Sequence 720 AA.

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:08:19 ; Search time 130.61 Seconds
(without alignments)
4.209 Million cell updates/sec

Title: PCT-US01-05825A-1
Perfect score: 46
Sequence: 1 GRVCVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_67:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	82.6	782	2	S18032 genome polyprotein
2	37	80.4	96	2	D83228 hypothetical prote
3	37	80.4	447	2	T34992 probable lipoprote
4	37	80.4	1691	1	D54689 protein-tyrosine-p
5	37	80.4	1894	2	C54689 protein-tyrosine-p
6	36	78.3	505	2	H83196 glycerol kinase PA
7	35	76.1	399	2	B82197 zona occludens t
8	35	76.1	399	2	A43864 zonula occludens t
9	35	76.1	680	2	PN0510 integrin beta-3 ch
10	35	76.1	723	2	PN0509 integrin beta-3 ch
11	34	73.9	92	2	D37057 epithelial cell gl
12	34	73.9	217	2	S66155 probable DNA-direc
13	34	73.9	593	2	S49525 glycoprotein G - s
14	34	73.9	753	2	B36268 platelet glycoprot
15	34	73.9	778	2	A60798 platelet glycoprot
16	34	73.9	788	2	A28547 platelet glycoprot
17	34	73.9	788	2	I77349 platelet glycoprot
18	34	73.9	1827	2	T34288 hypothetical prote
19	33	71.7	213	2	T32867 hypothetical prote
20	33	71.7	226	2	T20645 hypothetical prote
21	33	71.7	283	1	Q4ADG2 early B4 3K prote
22	33	71.7	298	2	T25874 hypothetical prote
23	33	71.7	440	2	T32673 hypothetical prote
24	33	71.7	470	2	A40697 63K sperm flagella
25	33	71.7	482	2	C72254 glycerol kinase -
26	33	71.7	494	2	F83199 probable carboxydr
27	33	71.7	508	2	S74537 anthranilate synth
28	33	71.7	517	2	A70793 probable glycerol
29	33	71.7	598	2	S39621 DNA-directed DNA p

30	71.7	600	2	B46642 DNA-directed DNA p
31	69.6	161	2	T21405 hypothetical prote
32	69.6	179	2	B64335 hypothetical prote
33	69.6	412	2	H65011 probable transamin
34	69.6	417	2	E83333 hypothetical prote
35	69.6	419	2	S42989 T48 protein - frui
36	69.6	447	1	A38561 alpha-1,3-mannosyl
37	69.6	500	2	T19525 hypothetical prote
38	69.6	508	1	B64204 glycerol kinase (E
39	69.6	570	1	JN0782 pyruvate decarboxy
40	69.6	672	2	T42186 conserved hypoxeth
41	69.6	698	2	T17261 hypothetical prote
42	69.6	1628	2	T43682 nucleoporin - fiss
43	69.6	2240	2	T37057 probable multi-dom
44	67.4	118	2	T35543 hypothetical prote
45	67.4	223	2	I41169 colicin A - Escher

ALIGNMENTS

RESULT 1
S18032 genome polyprotein - hepatitis C virus (isolate JK4) (fragment)
N:Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
C:Species: hepatitis C virus
A:Variety: isolate JK4
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
C:Accession: S18032
R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus
A:Reference number: S18029
A:Accession: S18032
A:Molecule type: genomic RNA
A:Residues: 1-782 <HON>
A:Cross-references: EMBL:X61594
A:Experimental source: isolate JK4
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructur
F:1-191/Product: core protein #status predicted <NAT1>
F:192-383/Product: envelope protein 1 #status predicted <NAT2>
F:384-733/Product: NS1/E2 protein #status predicted <NAT3>
F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <NAT4>

Query Match 82.6%; Score 38; DB 2; Length 782;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8
|||
Db 73 GRACVQPG 80

RESULT 2
D83228 hypothetical protein PA3338 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: D83228
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pc
A:Reference number: AB2950; MUID:20437337
A:Accession: D83228
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-96 <STO>
A:Cross-references: GB:AE004756; GB:AE004091; NID:99949466; PIDN:AAG06726.1; GSPDB:GN
A:Experimental source: strain PA01

C:Genetics:
A:Gene: PA3338

Query Match 80.4%; Score 37; DB 2; Length 96;
Best Local Similarity 62.5%; Pred. No. 4.1;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVOP: 8
|:|:|:|
Db 38 GKVCLEPG 45

RESULT 3

T34992
probable lipoprotein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T34992
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1998
A:Reference number: Z21550
A:Accession: T34992
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-447
A:Cross-references: EMBL:AL031182; PIDN:CAA20169.1; GSPDB:GN00070; SCOEDB:SC4A2.17C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC4A2.17C

Query Match 80.4%; Score 37; DB 2; Length 447;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVOP 7
|:|:|:|
Db 31 GRCLVOP 37

RESULT 4

protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type delta, splice form D precursor
N:Alternate names: MPTP delta type D
N:Contains: protein tyrosine phosphatase, receptor type delta, splice form A
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: D54689; A54689
R:Mizuno, K.; Hasegawa, K.; Katagiri, T.; Ogimoto, M.; Ichikawa, T.; Yakura, H.
Mol. Cell. Biol. 13, 5513-5523, 1993
A:Title: MPTP delta, a putative murine homolog of HPTP delta, is expressed in specialized
A:Reference number: A54689; MUID:93360986
A:Accession: D54689
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1691 <MIZ>
A:Experimental source: brain
A:Note: sequence inconsistent with nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:137486, NCBIP:136537)
A:Accession: A54689
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-398,799-1691 <MIZ>
A:Experimental source: brain
A:Note: sequence inconsistent with nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:136522, NCBIP:136524)
C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;
C:Keywords: alternative splicing; glycoprotein; phosphoprotein; phosphoric monoester hyd
F:42-95/Domain: immunoglobulin homology <IMM3>
F:114-196/Domain: fibronectin type III repeat homology <FN3A>
F:1075-1691/Domain: leukocyte common antigen cytosolic domain homology <LAC>

F:1449-1671/Domain: protein-tyrosine-phosphatase homology <PTP2>
F:1333/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1339/Binding site: substrate phosphate (Arg) #status predicted
F:1623/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1629/Binding site: substrate phosphate (Arg) #status predicted

Query Match 80.4%; Score 37; DB 1; Length 1691;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVOP 7
|:|:|:|
Db 1464 GRVCLQP 1470

RESULT 5

C54689
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type delta, splice form B precursor
N:Alternate names: MPTP delta type B/C
N:Contains: protein tyrosine phosphatase, receptor type delta, splice form C
C:Species: Mus musculus (house mouse)
C>Date: 25-Apr-1995 #sequence_revision 19-May-1995 #text_change 12-Feb-1999
C:Accession: C54689; B54689
R:Mizuno, K.; Hasegawa, K.; Katagiri, T.; Ogimoto, M.; Ichikawa, T.; Yakura, H.
Mol. Cell. Biol. 13, 5513-5523, 1993
A:Title: MPTP delta, a putative murine homolog of HPTP delta, is expressed in special
A:Reference number: A54689; MUID:93360986
A:Accession: C54689
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1894 <MIZ>
A:Experimental source: brain; splice form B
A:Note: sequence inconsistent with nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:137486, NCBIP:137487)
A:Accession: B54689
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-352,'H',354-535,'S',537-601,1002-1894 <MIZ>
A:Experimental source: brain; splice form C
A:Note: sequence inconsistent with nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:136527, NCBIP:136530)
C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology
C:Keywords: alternative splicing; glycoprotein; phosphoprotein; phosphoric monoester

F:45-107/Domain: immunoglobulin homology <IMM1>
F:245-239/Domain: immunoglobulin homology <IMM2>
F:317-399/Domain: fibronectin type III repeat homology <FN3A>
F:1278-1894/Domain: leukocyte common antigen cytosolic domain homology <LAC>
F:1652-1874/Domain: protein-tyrosine-phosphatase homology <PTP2>
F:1536/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1542/Binding site: substrate phosphate (Arg) #status predicted
F:1826/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1832/Binding site: substrate phosphate (Arg) #status predicted

Query Match 80.4%; Score 37; DB 2; Length 1894;
Best Local Similarity 85.7%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVOP 7
|:|:|:|
Db 1667 GRVCLQP 1673

RESULT 6

H83196
glycerol kinase PA3582 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83196
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrner, P.; Hickey, M.J.; L
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L

;; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337
A:Accession: H83196
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-505 <STO>
A:Cross-references: GB:AE004779; GB:AE004091; NID:g9949735; PIDN:AAG06970.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: glpK; PA3582
C:Superfamily: xylulokinase

Query Match 78.3%; Score 36; DB 2; Length 505;
Best Local Similarity 62.5%; Pred. No. 28;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8
II:III:II
Db 256 GOMCVERG 263

RESULT 7
B82197
zona occludens toxin VC1458 [Imported] - *Vibrio cholerae* (group O1 strain N16961)
C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: B82197
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.
L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MUID:20408833
A:Accession: B82197
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-399 <HEI>
A:Cross-references: GB:AE004224; GB:AE003852; NID:g9655952; PIDN:AAF94615.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1458
A:Map position: 1

Query Match 76.1%; Score 35; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8
II:III:II
Db 291 GRLCVQDG 298

RESULT 8
A43864
zonula occludens toxin - *Vibrio cholerae*
C:Species: *Vibrio cholerae*
C:Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C:Accession: A43864
R:Baudry, B.; Fasano, A.; Ketley, J.; Kaper, J.B.
Infect. Immun. 60, 428-434, 1992
A:Title: Cloning of a gene (zot) encoding a new toxin produced by *Vibrio cholerae*.
A:Reference number: A43864; MUID:92112300
A:Accession: A43864
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-399 <BAU>
A:Cross-references: GB:M83563; NID:g155314; PIDN:AAA27582.1; PID:g155315
A:Note: sequence extracted from NCBI backbone (NCBIN:77488, NCBIP:77491)

Query Match 76.1%; Score 35; DB 2; Length 723;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCVQPG 8
IIIIII
Db 548 VCVQPG 553

RESULT 11
D37057
epithelial cell glycoprotein Iiia - guinea pig (fragment)
C:Species: *Cavia porcellus* (guinea pig)
C:Date: 15-Feb-1991 #sequence_revision 15-Feb-1991 #text_change 23-Jul-1999
C:Accession: D37057
R:Sheppard, D.; Rozzo, C.; Starr, L.; Quaranta, V.; Erle, D.J.; Pytela, R.
J. Biol. Chem. 265, 11502-11507, 1990
A:Title: Complete amino acid sequence of a novel integrin beta subunit (beta6) ident1

Query Match 76.1%; Score 35; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8
II:III:II
Db 291 GRLCVQDG 298

RESULT 9
PN0510
Integrin beta-3 chain - mouse (fragment)
C:Species: *Mus musculus* (house mouse)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999
C:Accession: PN0510
R:Cieutat, A.M.; Rosa, J.P.; Letourneur, F.; Poncz, M.; Rifat, S.
Biochem. Biophys. Res. Commun. 193, 771-778, 1993
A:Title: A comparative analysis of cDNA-derived sequences for rat and mouse beta-3 in
A:Reference number: PN0509; MUID:93290675
A:Accession: PN0510
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-680 <CIE>
C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
C:Keywords: cell adhesion; duplication; heterodimer; membrane protein

Query Match 76.1%; Score 35; DB 2; Length 680;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCVQPG 8
IIIIII
Db 509 VCVQPG 514

RESULT 10
PN0509
Integrin beta-3 chain - rat (fragment)
C:Species: *Rattus norvegicus* (Norway rat)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999
C:Accession: PN0509
R:Cieutat, A.M.; Rosa, J.P.; Letourneur, F.; Poncz, M.; Rifat, S.
Biochem. Biophys. Res. Commun. 193, 771-778, 1993
A:Title: A comparative analysis of cDNA-derived sequences for rat and mouse beta-3 in
A:Reference number: PN0509; MUID:93290675
A:Accession: PN0509
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-723 <CIE>
C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
C:Keywords: cell adhesion; duplication; heterodimer; membrane protein

Query Match 76.1%; Score 35; DB 2; Length 723;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCVQPG 8
IIIIII
Db 548 VCVQPG 553

RESULT 11
D37057
epithelial cell glycoprotein Iiia - guinea pig (fragment)
C:Species: *Cavia porcellus* (guinea pig)
C:Date: 15-Feb-1991 #sequence_revision 15-Feb-1991 #text_change 23-Jul-1999
C:Accession: D37057
R:Sheppard, D.; Rozzo, C.; Starr, L.; Quaranta, V.; Erle, D.J.; Pytela, R.
J. Biol. Chem. 265, 11502-11507, 1990
A:Title: Complete amino acid sequence of a novel integrin beta subunit (beta6) ident1

A:Reference number: A37057; MUID:90307659

A:Accession: D37057

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-92 <SHE>

A:Cross-references: GB:J05522

C:Superfamily: integrin beta chain; laminin-type EGF-like homology

C:Keywords: glycoprotein

Query Match 73.9%; Score 34; DB 2; Length 92;

Best Local Similarity 83.3%; Pred. No. 15;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCVQPG 8

DB 82 VCIQPG 87

RESULT 12

S66155 Probable DNA-directed RNA polymerase (EC 2.7.7.6) - Pycnococcus provasolii (fragment)

N:Alternate names: T3/T7-like RNA polymerase

C:Species: Pycnococcus provasolii

C:Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 26-May-2000

C:Accession: S66155; S66160

R:Cermakian, N.; Ikeda, T.M.; Cedergren, R.; Gray, M.W.

submitted to the EMBL Data Library, August 1995

A:Reference number: S66152

A:Accession: S66155

A:Molecule type: DNA

A:Residues: 1-217 <CEW>

A:Cross-references: GB:U34286; NID:g1236335; PIDN:AAB01080.1; PID:g1236336

R:Cermakian, N.; Ikeda, T.M.; Cedergren, R.; Gray, M.W.

Nucleic Acids Res. 24, 648-654, 1996

A:Title: Sequences homologous to yeast mitochondrial and bacteriophage T3 and T7 RNA pol

A:Reference number: S66158; MUID:96177564

A:Accession: S66160

A:Molecule type: mRNA

A:Residues: 1-20; 33-66; 130-217 <CEW>

A:Cross-references: GB:U34286

A:Experimental source: CCMP 1203

C:Superfamily: phage T7 DNA-directed RNA polymerase

C:Keywords: nucleotidyltransferase; transcription

Query Match 73.9%; Score 34; DB 2; Length 217;

Best Local Similarity 85.7%; Pred. No. 31;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVQP 7

DB 142 GMVCVQP 148

RESULT 13

S49525

glycoprotein G - simian herpesvirus B

N:Alternate names: US4 protein homolog

C:Species: simian herpesvirus B

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Nov-1999

C:Accession: S49525

R:Slomka, M.J.; Brown, D.W.

submitted to the EMBL Data Library, October 1994

A:Description: Complete nucleotide sequence of simian herpes B virus glycoprotein G gene

A:Reference number: S49525

A:Accession: S49525

A:Molecule type: DNA

A:Residues: 1-593 <SLO>

A:Cross-references: EMBL:246268; NID:g560495; PIDN:CRAA86431.1; PID:g560496

A:Experimental source: isolate Cyno 2

C:Keywords: glycoprotein

Query Match 73.9%; Score 34; DB 2; Length 593;

Best Local Similarity 75.0%; Pred. No. 76;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8

DB 248 GRRCVSPG 255

RESULT 14

B36268 Platelet glycoprotein IIIa-II - human (fragment)

N:Alternate names: integrin beta-3' chain

C:Species: Homo sapiens (man)

C:Date: 18-Jan-1991 #sequence_revision 13-Jan-1993 #text_change 20-Aug-1999

C:Accession: B36268; A33907

R:Zimrin, A.B.; Gidwitz, S.; Lord, S.; Schwartz, E.; Bennett, J.S.; White II, G.C.; P

J. Biol. Chem. 265, 8590-8595, 1990

A:Title: The genomic organization of platelet glycoprotein IIIa.

A:Reference number: A36268; MUID:90256778

A:Accession: B36268

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-753 <ZIM>

A:Cross-references: GB:J05427

A>Note: the authors translated the codon GAT for residue 216 as Glu, GAT for residue

R:van Kuppevelt, T.H.M.S.M.; Languino, L.R.; Gallit, J.O.; Suzuki, S.; Ruoslahti, E.

Proc. Natl. Acad. Sci. U.S.A. 86, 5415-5418, 1989

A:Title: An alternative cytoplasmic domain of the integrin beta-3 subunit.

A:Reference number: A33907; MUID:89315807

A:Accession: A33907

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 580-753 <VAN>

A:Cross-references: GB:M25108; NID:g186502; PIDN:AAA36121.1; PID:g386833

C:Superfamily: integrin beta chain; laminin-type EGF-like homology

C:Keywords: cell adhesion; glycoprotein

Query Match 73.9%; Score 34; DB 2; Length 753;

Best Local Similarity 83.3%; Pred. No. 94;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCVQPG 8

DB 586 VCIQPG 591

RESULT 15

A60798

platelet glycoprotein IIIa beta chain (version 2) - human (fragment)

N:Alternate names: antigen CD61; integrin beta 3

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-Jul-1999

C:Accession: A60798; A36085; D32528; C32528; A36268

R:Rosa, J.P.; Bray, P.F.; Gayet, O.; Johnston, G.I.; Cook, R.G.; Jackson, K.W.; Shuma

Blood 72, 593-600, 1988

A:Title: Cloning of glycoprotein IIIa cDNA from human erythroleukemia cells and local

A:Reference number: A60798; MUID:88294314

A:Accession: A60798

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-778 <ROS>

R:Lanza, F.; Kieffer, N.; Phillips, D.R.; Fitzgerald, L.A.

J. Biol. Chem. 265, 18098-18103, 1990

A:Title: Characterization of the human platelet glycoprotein IIIa gene. Comparison wi

A:Reference number: A36085; MUID:91009291

A:Accession: A36085

A:Molecule type: DNA

A:Residues: 17-705, 'G', 707-778 <LAN>

A:Cross-references: GB:M57494

R:Hiraiwa, A.; Matsukage, A.; Shiku, H.; Takahashi, T.; Naito, K.; Yamada, K.

Blood 69, 560-564, 1987
A:Title: Purification and partial amino acid sequence of human platelet membrane glycoprotein IIb/IIIa
A:Reference number: A90731; MUID:87101510
A:Accession: D32528
A:Molecule type: protein
A:Residues: 208-224 <HI2>
A:Accession: C32528
A:Molecule type: protein
A:Residues: 429-433 <HIR>
R:Zimrin, A.B.; Gidwitz, S.; Lord, S.; Schwartz, E.; Bennett, J.S.; White II, G.C.; Pond
J. Biol. Chem. 265, 8590-8595, 1990
A:Title: The genomic organization of platelet glycoprotein IIIa.
A:Reference number: A36268; MUID:90256778
A:Accession: A36268
A:Molecule type: DNA
A:Residues: 18-705, 'G', 707-778 <ZIM>
A:Cross-references: GB:J05427
A:Note: the authors translated the codon GAT for residue 233 as Glu, GAT for residue 249
C:Genetics:
A:Gene: GDB:ITGB3
A:Cross-references: GDB:120013; OMIM:173470
A:Map position: 17q21.32-17q21.32
C:Superfamily: integrin beta chain; laminin-type EGF-like homology
C:Keywords: alternative splicing; glycoprotein

Query Match 73.9%; Score 34; DB 2; Length 778;
Best Local Similarity 83.3%; Pred. No. 97;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 VCQPG 8
||:|||
Db 603 VCIQPG 608

Search completed: June 13, 2001, 14:10:36
Job time: 137 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:08:20 ; Search time 74.44 Seconds
(without alignments)
3.681 Million cell updates/sec

Title: PCT-US01-05825A-1
Perfect score: 46
Sequence: 1 GRVCVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	36	78.3	505	1	GLPK_PSEAE
2	35	76.1	399	1	ZOT_VIBCH
3	35	76.1	787	1	ITB3_MOUSE
4	34	73.9	788	1	ITB3_HUMAN
5	33	71.7	283	1	EXXK_ADE02
6	33	71.7	298	1	VP26_ARATH
7	33	71.7	470	1	SP63_STRPU
8	33	71.7	496	1	GLK2_THEMA
9	33	71.7	508	1	TRPE_SYNY3
10	33	71.7	517	1	GLPK_MYCTU
11	33	71.7	589	1	DLL3_RAT
12	33	71.7	592	1	DLL3_MOUSE
13	33	71.7	598	1	DPO2_HUMAN
14	33	71.7	600	1	DPO2_MOUSE
15	33	71.7	618	1	DL13_HUMAN
16	33	71.7	636	1	P73_CERAE
17	33	71.7	637	1	PG20_AGRBL
18	32	69.6	30	1	PG20_AGRBL
19	32	69.6	179	1	V281_METJA
20	32	69.6	331	1	GL2M_ARATH
21	32	69.6	355	1	SUR6_MOUSE
22	32	69.6	412	1	YFD2_ECOLI
23	32	69.6	447	1	HTPX_METH
24	32	69.6	508	1	GLPK_MYCCE
25	32	69.6	570	1	DCPY_NEUCR
26	32	69.6	900	1	AXN1_HUMAN
27	32	69.6	992	1	AXN1_MOUSE
28	31	67.4	258	1	HTPX_METH
29	31	67.4	411	1	EP1G_TRYCR
30	31	67.4	445	1	GNT1_HUMAN
31	31	67.4	447	1	GNT1_MOUSE
32	31	67.4	447	1	GNT1_RAT
33	31	67.4	859	1	ENV_EIAV1

RESULT 1

ID	GLPK_PSEAE	STANDARD:	PRT:	505 AA.
AC	Q51390;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-OCT-2000 (Rel. 40, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE)			
DE	(GLYCEROKINASE) (GK).			
GN	GLPK OR PA3582.			
OS	Pseudomonas aeruginosa.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;			
OC	Pseudomonas.			
OX	NCBI_TaxID=287;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PA01;			
RX	MEDLINE=97286544; PubMed=9141691;			
RA	Schweizer H.P., Jump R., Po C.;			
RT	"Structure and gene-polypeptide relationships of the region encoding glycerol diffusion facilitator (glpF) and glycerol kinase (glpk) of Pseudomonas aeruginosa."			
RT	Microbiology 143:1287-1297(1997).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=PA01;			
RC	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen."			
RT	Nature 408:959-964(2000).			
CC	-1- FUNCTION: KEY ENZYME IN THE REGULATION OF GLYCEROL UPTAKE AND METABOLISM.			
CC	-1- CATALYTIC ACTIVITY: ATP + GLYCEROL = ADP + GLYCEROL 3-PHOSPHATE.			
CC	-1- PATHWAY: RATE-LIMITING STEP IN GLYCEROL UTILIZATION.			
CC	-1- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCOKINASE / GLYCEROKINASE / XYLOKINASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U49666; AAB57804.1; --			
DR	EMBL; AE004779; AAC06970.1; --			
DR	HSSP; P08859; IGLB.			
DR	InterPro; IPR000577; --			
DR	Pfam; PF00370; FGGY; 1.			

P22428 equine infe
P22429 equine infe
P11306 equine infe
P32541 equine infe
P16082 equine infe
P06751 equine infe
Q14028 homo sapien
P46582 caenorhabdi
P16615 h sarcoplas
P11607 s sarcoplas
P20647 o sarcoplas
O55143 m sarcoplas

ALIGNMENTS

DR PROSITE; PS00445; FGGY_KINASES.2; 1.
 DR PROSITE; PS00933; FGGY_KINASES.1; 1.
 KW Glycerol metabolism; Transferase; Kinase; ATP-binding.
 FT NP_BIND 158 170 ATP (PROBABLE).
 FT CONFLICT 75 75 H -> R (IN REF. 1).
 FT CONFLICT 104 104 A -> V (IN REF. 1).
 FT CONFLICT 109 109 C -> R (IN REF. 1).
 SQ SEQUENCE 505 AA; 55966 MW; CC63A9AF8ABCF752 CRC64;

Query Match 78.3%; Score 36; DB 1; Length 505;
 Best Local Similarity 62.5%; Pred. No. 10;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVOPG 8
 ::::|:|
 DB 256 GQMCVEEG 263

RESULT 2

ID ZOT_VIBCH STANDARD; PRT; 399 AA.
 AC P38442; Q9L706; Q9R3V6;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ZONA OCCLUDENS TOXIN (ZONULAR OCCLUDENS TOXIN).
 GN ZOT OR VC1458.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_Taxid=566;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-CLASSICAL INABA 569B;
 RX MEDLINE-92112300; PubMed-1730472;
 RA Baudry B., Fasano A., Ketley J., Kaper J.B.;
 RT "Cloning of a gene (zot) encoding a new toxin produced by Vibrio
 cholerae.";
 RL Infect. Immun. 60:428-434(1992).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-KNIH002;
 RA Shin H.J., Park Y.C., Kim Y.C.;
 RT "Cloning and nucleotide sequence analysis of the virulence gene
 cassette from Vibrio cholerae KNIH002 isolated in Korea.";
 RL Misalimurhag Hoiji 35:205-210(1999).
 [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN-OL39-TOR OGAWA;
 RA Zhi-Yong H., Wei-Jie Z., Xiang-Fu W.;
 RT "Cloning and Expression of zot Gene from Vibrio cholerae.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN-EL TOR 86015 / SEROTYPE O1;
 RA Kan B., Liu Y.Q., Qi G.M., Gao S.Y.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 [5]
 RN SEQUENCE FROM N.A.
 RC STRAIN-EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE-20406833; PubMed-10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae.";
 RL Nature 406:477-483(2000).
 [6]
 RN CHARACTERIZATION.
 RP

RX MEDLINE-91271365; PubMed-2052603;
 RA Fasano A., Baudry B., Pumphlin D.W., Wasserman S.S., Tall B.D.,
 RA Ketley J.M., Kaper J.B.;
 RT "Vibrio cholerae produces a second enterotoxin, which affects
 intestinal tight junctions.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5242-5246(1991).
 CC -!- FUNCTION: INCREASES THE PERMEABILITY OF THE SMALL INTESTINE MUCOSA
 BY AFFECTING THE STRUCTURE OF INTERCELLULAR TIGHT JUNCTIONS
 (ZONULA OCCLUDENS).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M83563; AAA27582.1; -;
 DR EMBL; AF175708; AAD51358.1; -;
 DR EMBL; AF123049; AAD26854.1; -;
 DR EMBL; AF220606; AAF29547.1; -;
 DR EMBL; AE004224; AAF94615.1; -;
 DR EIR; A43864; A43864.
 DR TIGR; VC1458; -;
 KW Enterotoxin; Toxin.
 FT VARIANT 45 45 M -> I (IN STRAIN 569B).
 FT VARIANT 100 100 V -> A (IN STRAIN 569B AND 86015).
 FT VARIANT 272 272 V -> A (IN STRAIN 569B).
 FT VARIANT 281 281 V -> A (IN STRAIN 569B).
 FT VARIANT 349 349 A -> S (IN STRAIN 86015).
 FT VARIANT 381 381 K -> R (IN STRAIN 86015).
 FT CONFLICT 386 399 IKTENDKGLNSIF -> VKKEEESIKSFL (IN REF.
 4).
 FT SEQUENCE 399 AA; 44903 MW; 3C7424B758176774 CRC64;
 SQ
 Query Match 76.1%; Score 35; DB 1; Length 399;
 Best Local Similarity 75.0%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GRVCVOPG 8
 ::::|:|
 DB 291 GRUCVODG 298
 RESULT 3
 ITB3_MOUSE STANDARD; PRT; 787 AA.
 ID ITB3_MOUSE
 AC O54890;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PLATELET MEMBRANE GLYCOPROTEIN IIIA PRECURSOR (GPIIIA) (INTEGRIN BETA-
 3) (CD61).
 GN ITGB3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C3H/HEN;
 RA McHugh K.P., Teitelbaum S.L., Kitazawa S., Ross F.P.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: PLATELET MEMBRANE GLYCOPROTEIN IIB-IIIA IS THE PLATELET
 RECEPTOR FOR ADHESIVE PROTEINS. IT BINDS FIBRINOGEN, FIBRONECTIN,
 CC VITRONECTIN AND VWF. INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN
 CC THEIR LIGAND.
 CC -!- SUBUNIT: DIMER OF AN ALPHA AND BETA SUBUNIT. BETA-3 IS EITHER
 CC ASSOCIATED WITH ALPHA-IIB (GPIIB) OR WITH ALPHA-V (VITRONECTIN
 CC RECEPTOR).
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS.
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF026509; AAB94086.1; -
 CC MGD; MGI:96612; Itgb3.
 CC InterPro: IPR000561; -
 CC InterPro: IPR001169; -
 CC InterPro: IPR002369; -
 CC Pfam: PF00362; Integrin_B; 1.
 CC PRINTS; PRO1186; INTEGRIN.
 CC PROSITE; PS00243; INTEGRIN_BETA; 3.
 CC PROSITE; PS00022; EGF_1; UNKNOWN_2.
 CC PROSITE; PS01186; EGF_2; UNKNOWN_1.
 CC Integrin; Cell adhesion; Transmembrane; Glycoprotein; Repeat;
 CC Extracellular matrix; Cytoskeleton; Signal.
 CC SIGNAL 1 25 POTENTIAL.
 CC CHAIN 1 25 PLATELET MEMBRANE GLYCOPROTEIN IIIA.
 CC DOMAIN 26 717 EXTRACELLULAR.
 CC TRANSMEM 718 740 POTENTIAL.
 CC DOMAIN 741 787 CYTOPLASMIC.
 CC DOMAIN 462 628 CYSTEINE-RICH REPEATS.
 CC REPEAT 462 510 I.
 CC REPEAT 511 552 II.
 CC REPEAT 553 591 III.
 CC REPEAT 592 628 IV.
 CC DISULFID 30 460 BY SIMILARITY.
 CC DISULFID 38 48 BY SIMILARITY.
 CC DISULFID 41 74 BY SIMILARITY.
 CC DISULFID 51 63 BY SIMILARITY.
 CC DISULFID 202 209 BY SIMILARITY.
 CC DISULFID 257 298 BY SIMILARITY.
 CC DISULFID 399 411 BY SIMILARITY.
 CC DISULFID 431 680 BY SIMILARITY.
 CC DISULFID 458 462 BY SIMILARITY.
 CC DISULFID 473 485 BY SIMILARITY.
 CC DISULFID 482 520 BY SIMILARITY.
 CC DISULFID 487 496 BY SIMILARITY.
 CC DISULFID 498 511 BY SIMILARITY.
 CC DISULFID 526 531 BY SIMILARITY.
 CC DISULFID 528 561 BY SIMILARITY.
 CC DISULFID 533 546 BY SIMILARITY.
 CC DISULFID 548 553 BY SIMILARITY.
 CC DISULFID 567 572 BY SIMILARITY.
 CC DISULFID 569 600 BY SIMILARITY.
 CC DISULFID 574 583 BY SIMILARITY.
 CC DISULFID 585 592 BY SIMILARITY.
 CC DISULFID 606 611 BY SIMILARITY.
 CC DISULFID 608 656 BY SIMILARITY.
 CC DISULFID 613 623 BY SIMILARITY.
 CC DISULFID 626 629 BY SIMILARITY.
 CC DISULFID 633 642 BY SIMILARITY.
 CC DISULFID 639 712 BY SIMILARITY.
 CC DISULFID 660 688 BY SIMILARITY.
 CC CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 477 477 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 584 584 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SEQUENCE 787 AA; 86694 MW; B1570599ABC438A3 CRC64;

Query Match

76.1%; Score 35; DB 1; Length 787;

Best Local Similarity 100.0%; Pred. No. 24;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCQVPG 8
 Db 612 VCQVPG 617
 RESULT 4
 ITB3_HUMAN
 ID ITB3_HUMAN STANDARD; PRT; 788 AA.
 AC P05106; Q13413; Q16499;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PLATELET MEMBRANE GLYCOPROTEIN IIIA PRECURSOR (GPIIIA) (INTEGRIN BETA-
 DE 3) (CD61).
 DE ITGB3 OR GP3A.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87165991; PubMed=3494014;
 RA Fitzgerald L.A., Steiner B., Rall S.C. Jr., Lo S., Phillips D.R.;
 RT "Protein sequence of endothelial glycoprotein IIIa derived from a
 RT cDNA clone. Identity with platelet glycoprotein IIIa and similarity
 RT to 'integrin'.";
 RL J. Biol. Chem. 262:3936-3939(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90265363; PubMed=2345548;
 RA Frachet P., Uzan G., Thevenon D., Denarier E., Prandini M.H.,
 RA Marguerie G.;
 RT "GPIIb and GPIIIa amino acid sequences deduced from human
 RT megakaryocyte cDNAs.";
 RL Mol. Biol. Rep. 14:27-33(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88213696; PubMed=2452834;
 RA Zimrin A.B., Eisman R., Villaire G., Schwartz E., Bennett J.S.,
 RA Poncz M.;
 RT "Structure of platelet glycoprotein IIIa. A common subunit for two
 RT different membrane receptors.";
 RL J. Clin. Invest. 81:1470-1475(1988).
 RN [4]
 RP SEQUENCE OF 27-788 FROM N.A.
 RX MEDLINE=91009291; PubMed=2145280;
 RA Lanza F., Kieffer N., Phillips D.R., Fitzgerald L.A.;
 RT "Characterization of the human platelet glycoprotein IIIa gene.
 RT Comparison with the fibronectin receptor beta-subunit gene.";
 RL J. Biol. Chem. 265:18098-18103(1990).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90256778; PubMed=2341395;
 RA Zimrin A.B., Gidwitz S., Lord S., Schwartz E., Bennett J.S.,
 RA White G.C. II, Poncz M.;
 RT "The genomic organization of platelet glycoprotein IIIa.";
 RL J. Biol. Chem. 265:8590-8595(1990).
 RN [6]
 RP SEQUENCE OF 1-26 FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=94129007; PubMed=8298129;
 RA Villa-Garcia M., Li L., Riely G., Bray P.F.;
 RT "Isolation and characterization of a TATA-less promoter for the human
 RT beta 3 integrin gene.";
 RL Blood 83:668-676(1994).
 RN [7]
 RP SEQUENCE OF 122-204 FROM N.A.
 RX MEDLINE=93002753; PubMed=1382574;
 RA Jiang W.M., Jenkins D., Yuan Q., Leung E., Choo K.H., Watson J.D.,
 RA Krissansen G.W.;
 RT "The gene organization of the human beta 7 subunit, the common beta
 RT subunit of the leukocyte integrins HML-1 and LPAM-1.";

RL Int. Immunol. 4:1031-1040(1992).
 RN [8]
 RP SEQUENCE OF 218-234 AND 439-443.
 RX MEDLINE=87101510; PubMed=3801670;
 RA Hiraiva A., Matsukage A., Shiku H., Takahashi T., Naito K., Yamada K.;
 RT "Purification and partial amino acid sequence of human platelet
 RL membrane glycoproteins IIB and IIIA";
 RN Blood 69:560-564(1987).
 RP [9]
 RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
 RX MEDLINE=91158732; PubMed=2001252;
 RA Calvete J.J., Henschen A., Gonzalez-Rodriguez J.;
 RT "Assignment of disulphide bonds in human platelet GPIIb. A
 RL biochem. J. 274:63-71(1991).
 RN [10]
 RP VARIANT HPA-1 (PL(A)).
 RX MEDLINE=89214713; PubMed=2565345;
 RA Newman P.J., Derbes R.S., Aster R.H.;
 RT "The human platelet alloantigens, PLAI and PL2, are associated with
 RL a leucine33/proline33 amino acid polymorphism in membrane
 glycoprotein IIb, and are distinguishable by DNA typing.";
 RN J. Clin. Invest. 83:1778-1781(1989).
 RP [11]
 RP VARIANT HPA-4 (PEN).
 RX MEDLINE=93053444; PubMed=1430225;
 RA Wang R., Furihata K., McFarland J.G., Friedman K., Aster R.H.,
 RA Newman P.J.;
 RT "An amino acid polymorphism within the RGD binding domain of platelet
 RL membrane glycoprotein IIb is responsible for the formation of the
 PENA/PENb alloantigen system.";
 RN J. Clin. Invest. 90:2038-2043(1992).
 RP [12]
 RP VARIANT MO-1.
 RX MEDLINE=93112977; PubMed=8093349;
 RA Kullpers R.W.A.M., Simsek S., Faber N.M., Goldschmeding R.,
 RA van Wermkerken R.K.V., von dem Borne A.E.G.K.;
 RT "Single point mutation in human glycoprotein IIb is associated with
 RL a new platelet-specific alloantigen (MO) involved in neonatal
 alloimmune thrombocytopenia.";
 RN Blood 81:70-76(1993).
 RP [13]
 RP VARIANT CA/TU.
 RX MEDLINE=94060373; PubMed=7694683;
 RA Wang R., McFarland J.G., Kekomaki R., Newman P.J.;
 RT "Amino acid 489 is encoded by a mutational 'hot spot' on the beta 3
 RL integrin chain: the CA/TU human platelet alloantigen system.";
 RN Blood 82:3386-3391(1993).
 RP [14]
 RP VARIANT SR(A).
 RX MEDLINE=94179229; PubMed=8132570;
 RA Santoso S., Kalb R., Kroll H., Walika M., Kiesel V.,
 RA Mueller-Eckhardt C., Newman P.J.;
 RT "A point mutation leads to an unpaired cysteine residue and a
 RL molecular weight polymorphism of a functional platelet beta 3 integrin
 subunit. The Sra alloantigen system of GPIIb.";
 RN J. Biol. Chem. 269:8439-8444(1994).
 RP [15]
 RP VARIANT TVR-145.
 RX MEDLINE=90364410; PubMed=2392682;
 RA Loftus J.C., O'Toole T.E., Plow E.F., Glass A., Frelinger A.L. III,
 RA Ginsberg M.H.;
 RT "A beta 3 integrin mutation abolishes ligand binding and alters
 RL divalent cation-dependent conformation.";
 RN Science 249:915-918(1990).
 RP [16]
 RP VARIANT SER-240.
 RX MEDLINE=92156115; PubMed=1371279;
 RA Bajt M.L., Ginsberg M.H., Frelinger A.L. III, Berndt M.C.,
 RA Loftus J.C.;
 RT "A spontaneous mutation of integrin alpha IIB beta 3 (platelet
 RL glycoprotein IIb-IIIA) helps define a ligand binding site.";
 RN J. Biol. Chem. 267:3789-3794(1992).

RN [17]
 RP VARIANT TRP-240.
 RX MEDLINE=92291320; PubMed=1602006;
 RA Lanza F., Stierle A., Fournier D., Morales M., Nurdan A.T.,
 RA Cazenave J.-P.;
 RT "A new variant of Glanzmann's thrombasthenia (Strasbourg I).
 RL Platelets with functionally defective glycoprotein IIb-IIIa complexes
 and a glycoprotein IIa 214Arg->214Trp mutation.";
 RN J. Clin. Invest. 89:1995-2004(1992).
 RP [18]
 RP VARIANT STRASBOURG-1.
 RX MEDLINE=93066201; PubMed=1438206;
 RA Chen Y.-P., Djaffar I., Pidard D., Steiner B., Cieutat A.-M.,
 RA Caen J.P., Rosa J.-P.;
 RT "Ser-752->Pro mutation in the cytoplasmic domain of integrin beta 3
 RL subunit and defective activation of platelet integrin alpha IIB beta
 3 (glycoprotein IIb-IIIa) in a variant of Glanzmann thrombasthenia.";
 RN Proc. Natl. Acad. Sci. U.S.A. 89:10169-10173(1992).
 RP [19]
 RP REVIEW ON GTA VARIANTS.
 RX MEDLINE=95184171; PubMed=7878622;
 RA Bray P.F.;
 RT "Inherited diseases of platelet glycoproteins: considerations for
 RL rapid molecular characterization.";
 RN Thromb. Haemost. 72:492-502(1994).
 RP [20]
 RP VARIANTS GTA PRO-306; PHE-586 AND SER-598.
 RX MEDLINE=99008899; PubMed=9790984;
 RA Ambo H., Kamata T., Handa M., Taki M., Kuwajima M., Kawai Y., Oda A.,
 RA Murata M., Takada Y., Watanabe K., Ikeda Y.;
 RT "Three novel integrin beta3 subunit missense mutations (H280P, C560F,
 RL and G579S) in thrombasthenia, including one (H280P) prevalent in
 Japanese patients.";
 RN Biochem. Biophys. Res. Commun. 251:763-768(1998).
 CC -1- FUNCTION: PLATELET MEMBRANE GLYCOPROTEIN IIB-IIIA IS THE PLATELET
 RECEPTOR FOR ADHESIVE PROTEINS. IT BINDS FIBRINOGEN, FIBRONECTIN,
 CC VITRONECTIN AND VWF. INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN
 CC THEIR LIGAND.
 CC -1- SUBUNIT: DIMER OF AN ALPHA AND BETA SUBUNIT. BETA-3 IS EITHER
 CC ASSOCIATED WITH ALPHA-IIB (GPIIB) OR WITH ALPHA-V (VITRONECTIN
 CC RECEPTOR).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: THE DIFFERENT FORMS OF BETA-3 SUBUNIT MAY
 CC ARISE BY ALTERNATIVE SPLICING OF PRIMARY MRNA TRANSCRIPTS.
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS.
 CC -1- POLYMORPHISM: POSITION 59 IS ASSOCIATED WITH PLATELET-SPECIFIC
 CC ALLOANTIGEN HPA-1 (2W OR PL(A)). HPA-1A/PL(A1) HAS LEU-59 AND
 CC HPA-1B/PL(A2) HAS PRO-59.
 CC -1- POLYMORPHISM: POSITION 169 IS ASSOCIATED WITH PLATELET-SPECIFIC
 CC ALLOANTIGEN HPA-4 (PEN OR YUK). HPA-4A/PEN(A)/YUK(A) HAS ARG-169
 CC AND HPA-4B/PEN(B)/YUK(B) HAS GLN-169. HPA-4B IS INVOLVED IN
 CC NEONATAL ALLOIMMUNE THROMBOCYTOPENIA (NATP).
 CC -1- POLYMORPHISM: POSITION 433 IS ASSOCIATED WITH PLATELET-SPECIFIC
 CC ALLOANTIGEN MO. MO(-) HAS PRO-433 AND MO(+) HAS ALA-433. MO(+) IS
 CC INVOLVED IN NATP.
 CC -1- POLYMORPHISM: POSITION 515 IS ASSOCIATED WITH PLATELET-SPECIFIC
 CC ALLOANTIGEN CA (TU). CA(-)/TU(-) HAS ARG-515 AND CA(1)/TU(+) HAS
 CC GLN-515. CA(+) IS INVOLVED IN NATP.
 CC -1- DISEASE: DEFECTS IN ITGB3 ARE ONE OF THE CAUSE OF GLANZMANN
 CC THROMBASTHENIA (GTA), AN AUTOSOMAL RECESSIVE DISORDER WHICH IS
 CC THE MOST COMMON INHERITED DISEASE OF PLATELETS. GTA IS

Query Match 73.9%; Score 34; DB 1; Length 788;
 Best Local Similarity 83.3%; Pred. No. 38;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 VCVQPG 8
 Db 613 VCIQPG 618

RESULT 5

```

EXXK_ADE02
ID EXXK_ADE02 STANDARD; PRT; 283 AA.
AC P03242;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PROBABLE EARLY E4 PROTEINS.
OS Human adenovirus type 2.
OX Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OC NCBI_TaxID=10515;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82059444; PubMed=6985482;
RA Herisse J., Rigolet M., Dupont de Dinechin S., Galibert F.;
RT "Nucleotide sequence of adenovirus 2 DNA fragment encoding for the
RL carboxylic region of the fiber protein and the entire E4 region.";
RL Nucleic Acids Res. 9:4023-4042(1981).
CC -1- MISCELLANEOUS: THESE PROBABLE PROTEINS AND THE INTRONS IN THE
CC CODING REGIONS WERE ASSIGNED BY CORRELATING EM DATA, S1 DIGESTION
CC STUDIES, AND THE CONSENSUS SEQUENCES FOR INTRON SPLICING.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J01917; ; NOT_ANNOTATED_CDS.
DR PIR; A03808; Q4ADG2.
DR InterPro; IPR001428; .
DR Pfam; PF00692; dufpase; 1.
KW Early protein.
FT CHAIN 1 283 EARLY 31 KDA PROTEIN.
FT CHAIN 1 255 EARLY 28 KDA PROTEIN.
FT CHAIN 1 121 EARLY 14 KDA PROTEIN 2 + C-TER VROASNV.
FT CHAIN 122 255 EARLY 14 KDA PROTEIN + N-TER MQ.
FT CHAIN 223 255 SECOND PART OF EARLY PROTEIN 17K.
FT CHAIN 223 283 SECOND PART OF EARLY PROTEIN 20K.
FT CHAIN 160 255 SECOND PART OF EARLY PROTEIN 24K.
FT CHAIN 160 283 SECOND PART OF EARLY PROTEIN 27K.
FT DOMAIN 1 121 COMMON PART OF EARLY PROTEINS 17K, 20K,
FT 24K AND 27K.
SQ SEQUENCE 283 AA; 31846 MW; 26382A77A65C14DB CRC64;

Query Match 71.7%; Score 33; DB 1; Length 283;
Best Local Similarity 71.4%; Pred. No. 23;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVCVQPG 8
Db |||: ||
52 RVCIPPG 58

RESULT 6
VP26_ARATH STANDARD; PRT; 298 AA.
ID VP26_ARATH
AC Q9T091;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VPS26 PROTEIN HOMOLOG (FRAGMENT).
GN A74G27690 OR T29A15.180.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;

MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansorge W., Brandt P., Grivell L., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Gijmonprez B., Chuang Y.-J., Vandenbusche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Berneriser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA Pettett A., Rajadream M.-A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloecker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Feiber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chefdr F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacou D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bieleke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegler L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latrelle P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Manx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana."
RL Nature 402:769-777(1999).
CC -1- SIMILARITY: BELONGS TO THE VPS26 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL035602; CAB38281.1; .
DR EMBL; AL161571; CAB81419.1; .
KW Hypothetical protein; Transport; Protein transport.
FT NON_TER 1
SQ SEQUENCE 298 AA; 34462 MW; 576AB065C2052DC0 CRC64;

Query Match 71.7%; Score 33; DB 1; Length 298;
Best Local Similarity 57.1%; Pred. No. 25;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVQP 7
Db |::|:|
48 GKVCIEP 54

```

```

RESULT 7
SP63_STRPU STANDARD; PRT; 470 AA.
AC Q07929; P98117;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 63 KDA SPERM FLAGELLAR MEMBRANE PROTEIN PRECURSOR.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoides; Euechinozoa; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93286182; PubMed=8509450;
RA Mendoza L.M., Nishio D., Vacquier V.D.;
RT "A GPI-anchored sea urchin sperm membrane protein containing EGF
RL domains is related to human uromodulin."
RL J. Cell Biol. 121:1291-1297(1993).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE FLAGELLAR MEMBRANE BY A
    GPI-ANCHOR.
CC -1- TISSUE SPECIFICITY: SPERM.
CC -1- SIMILARITY: CONTAINS 1 SEA DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M9584; AAA30029.1;
DR InterPro: IPR000082;
DR InterPro: IPR000152;
DR InterPro: IPR000561;
DR InterPro: IPR001881;
DR Pfam: PF01390; SEA; 1.
DR PROSITE: PS00010; ASX-HYDROXYL; 1.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA; 1.
KW Repeat: EGF-like domain; Signal; Membrane; Sperm; GPI-anchor;
KW Glycoprotein.
FT SIGNAL 1 25
FT CHAIN 26 446
FT PROPEP 447 470
FT DOMAIN 41 80
FT DOMAIN ? ?
FT DOMAIN 202 250
FT DOMAIN 249 292
FT DISULFID 45 57
FT DISULFID 50 66
FT DISULFID 68 79
FT DISULFID 206 220
FT DISULFID 214 229
FT DISULFID 231 249
FT DISULFID 253 265
FT DISULFID 258 277
FT DISULFID 279 291
FT CARBOHYD 78 78
FT CARBOHYD 170 170
FT CARBOHYD 219 219
FT CARBOHYD 322 322
FT CARBOHYD 446 446
FT LIPID 470 AA; 51101 MW; 4D42BC147D0DID6 CRC64;
SQ SEQUENCE 470 AA; 51101 MW; 4D42BC147D0DID6 CRC64;

```

```

Query Match 71.7%; Score 33; DB 1; Length 470;
Best Local Similarity 71.4%; Pred. No. 37;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 GRVCVQVP 7
DB 246 GRVCIAP 252
ID GLK2_THEME STANDARD; PRT; 496 AA.
AC Q9X1E4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GLYCEROL KINASE 2 (EC 2.7.1.30) (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE 2)
DE (GLYCEROKINASE 2) (GK 2).
GN GLPK2 OR TM1430.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
CC -1- FUNCTION: KEY ENZYME IN THE REGULATION OF GLYCEROL UPTAKE AND
    METABOLISM.
CC -1- CATALYTIC ACTIVITY: ATP + GLYCEROL = ADP + GLYCEROL 3-PHOSPHATE.
CC -1- PATHWAY: RATE-LIMITING STEP IN GLYCEROL UTILIZATION.
CC -1- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCONOKINASE /
    GLYCEROKINASE / XYLULOSE KINASE FAMILY.
CC -1- CAUTION: A STOP CODON IN POSITION 483 WAS TRANSLATED AS TRP TO
    RESTORE THE SIMILARITY WITH THE C-TERMINAL REGION OF OTHERS GLPK
    HOMOLOGS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001795; AAD36500.1; ALT_SEQ.
DR TIGR: TM1430;
DR InterPro: IPR000577;
DR Pfam: PF00370; FGGY; 1.
DR PROSITE: PS00933; FGGY_KINASES_1; 1.
DR PROSITE: PS00445; FGGY_KINASES_2; 1.
KW Glycerol metabolism; transferase; Kinase; ATP-binding.
FT NP_BIND 152 164 ATP (PROBABLE).
FT SEQUENCE 496 AA; 55576 MW; 7FA72A4CBD29E701 CRC64;
SQ SEQUENCE 496 AA; 55576 MW; 7FA72A4CBD29E701 CRC64;

```

```

Query Match 71.7%; Score 33; DB 1; Length 496;
Best Local Similarity 62.5%; Pred. No. 39;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GRVCVQVP 8
DB 249 GQACFQPC 256
ID TRPE_SYNV3 STANDARD; PRT; 508 AA.
AC P20170; P20168;

```

```

RESULT 9
TRPE_SYNV3
ID TRPE_SYNV3 STANDARD; PRT; 508 AA.
AC P20170; P20168;

```

DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROBABLE ANTHRANILATE SYNTHASE COMPONENT I (EC 4.1.3.27).
 GN TRPE OR SLR0738.
 OS *Synechocystis* sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; *Synechocystis*.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
 RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
 RA Yamada M., Yasuda M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT *Synechocystis* sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 RN [2]
 RP SEQUENCE OF 1-408 FROM N.A.
 RX MEDLINE=89034300; PubMed=3141423;
 RA Reilly P., Hulmes J.D., Pan Y.C.E., Nelson N.;
 RT "Molecular cloning and sequencing of the psad gene encoding subunit
 RT II of photosystem I from the cyanobacterium, *Synechocystis* sp. PCC
 RT 6803.";
 RL J. Biol. Chem. 263:17658-17662(1988).
 CC -|- CATALYTIC ACTIVITY: CHORISWATE + L-GLUTAMINE = ANTHRANILATE +
 CC PYRUVATE + L-GLUTAMATE.
 CC -|- PATHWAY: FIRST STEP IN BIOSYNTHESIS OF TRYPTOPHAN.
 CC -|- SUBUNIT: Tetramer of two components I and two components II (by
 CC similarity).
 CC -|- MISCELLANEOUS: COMPONENT I CATALYZES THE FORMATION OF ANTHRANILATE
 CC USING AMMONIA RATHER THAN GLUTAMINE, WHEREAS COMPONENT II PROVIDES
 CC GLUTAMINE AMIDOTRANSFERASE ACTIVITY.
 CC -|- SIMILARITY: BELONGS TO THE ANTHRANILATE SYNTHASE COMPONENT I
 CC FAMILY.
 CC -|- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
 CC FRAMESHIFT IN POSITION 182.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D90899; BAA16689.1; -;
 DR EMBL; J04195; AAA88626.1; ALT_FRAME.
 DR EMBL; J04195; AAA88627.1; ALT_FRAME.
 DR PIR; B32124; B32124.
 DR PIR; C32124; C32124.
 DR InterPro: IPR000350; -;
 DR Pfam: PF00425; chorismate_bind; 1.
 DR PRINTS: PR00095; ANTSNTHASE1.
 KW Tryptophan biosynthesis; Lyase.
 SQ SEQUENCE 508 AA; 57147 MW; A3308E9A7B8A1FC1 CRC64;

Query Match 71.7%; Score 33; DB 1; Length 508;
 Best Local Similarity 100.08; Pred. No. 40;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVQ 6

Db 363 GRVCVQ 368

RESULT 10

GLPK_MYCTU

ID GLPK_MYCTU STANDARD; PRT; 517 AA.

AC 069664;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE)
 DE (GLYCEROKINASE) (GK).
 DE GLPK OR RV3696C OR MTV025.044C.
 GN Mycobacterium tuberculosis.
 OS Mycobacteriaceae; Actinobacteria; Mycobacterium.
 OC Bacteria; Firmicutes; Actinobacteriia; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver K., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton J., Squares R., Squares S.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 CC -|- FUNCTION: KEY ENZYME IN THE REGULATION OF GLYCEROL UPTAKE AND
 CC METABOLISM.
 CC -|- CATALYTIC ACTIVITY: ATP + GLYCEROL = ADP + GLYCEROL 3-PHOSPHATE.
 CC -|- PATHWAY: RATE-LIMITING STEP IN GLYCEROL UTILIZATION.
 CC -|- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCONOKINASE /
 CC GLYCEROKINASE / XYLULOKINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL022121; CAAL8018.1; -;
 DR HSSP; P08859; 1GLB.
 DR TuberculList; RV3696C; -;
 DR InterPro: IPR000577; -;
 DR Pfam: PF00370; FGXY; 1.
 DR PROSITE: PS00445; FGXY_KINASES_2; 1.
 DR PROSITE: PS00933; FGXY_KINASES_1; 1.
 DR Glycylol metabolism; Transferase; Kinase; ATP-binding.
 FT NP_BIND 165 177
 FT ATP (PROBABLE).
 SQ SEQUENCE 517 AA; 55859 MW; 7E3F30DCF2A1E63E CRC64;
 Query Match 71.7%; Score 33; DB 1; Length 517;
 Best Local Similarity 62.5%; Pred. No. 40;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8

Db 268 GOVCLAPG 275

RESULT 11

DL33_RAT

ID DL33_RAT STANDARD; PRT; 589 AA.

AC O88671;

DT 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE DELTA-LIKE PROTEIN 3 PRECURSOR (DROSOPHILA DELTA HOMOLOG 3).

GN DLL3.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RA SEQUENCE FROM N.A.
RP Boulter J., Greenfield A., Weinmaster G.;
RT "Rattus norvegicus mRNA for Delta 3: a putative ligand for Notch.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INHIBITS PRIMARY NEUROGENESIS. MAY BE REQUIRED TO DIVERT
CC NEURONS ALONG A SPECIFIC DIFFERENTIATION PATHWAY. PLAY A ROLE IN
CC THE FORMATION OF SOMITE BOUNDARIES DURING SEGMENTATION OF THE
CC PARAXIAL MESODERM (BY SIMILARITY).
CC -1- SUBUNIT: CAN BIND AND ACTIVATE NOTCH-1 OR ANOTHER NOTCH RECEPTOR
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- DOMAIN: THE DELTA-SERRATE-LAG2 (DSL) DOMAIN IS REQUIRED FOR
CC BINDING TO THE NOTCH RECEPTOR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF084576; AAC33303.1; -;
DR HSP; P00740; 11XA;
DR InterPro; IPR000561; -;
DR InterPro; IPR000742; -;
DR InterPro; IPR001438; -;
DR Pfam; PF00008; EGF; 5;
DR PROSITE; PS00022; EGF.1; 6;
DR PROSITE; PS01186; EGF.2; 5;
KW Signal; EGF-like domain; Repeat; Transmembrane; Developmental protein;
KW Differentiation.
FT SIGNAL 1 32 BY SIMILARITY.
FT CHAIN 33 589 DELTA-LIKE PROTEIN 3.
FT DOMAIN 33 494 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 495 515 POTENTIAL.
FT DOMAIN 516 589 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 174 213 DELTA-SERRATE-LAG2.
FT DOMAIN 239 250 EGF-LIKE 1.
FT DOMAIN 300 311 EGF-LIKE 2.
FT DOMAIN 341 352 EGF-LIKE 3.
FT DOMAIN 379 390 EGF-LIKE 4.
FT DOMAIN 417 428 EGF-LIKE 5.
FT DOMAIN 455 466 EGF-LIKE 6.
FT DOMAIN 503 511 POLY-ALA.
FT DISULFID 241 250 BY SIMILARITY.
FT DISULFID 285 300 BY SIMILARITY.
FT DISULFID 302 311 BY SIMILARITY.
FT DISULFID 329 341 BY SIMILARITY.
FT DISULFID 343 352 BY SIMILARITY.
FT DISULFID 359 370 BY SIMILARITY.
FT DISULFID 364 379 BY SIMILARITY.
FT DISULFID 381 390 BY SIMILARITY.
FT DISULFID 397 408 BY SIMILARITY.
FT DISULFID 402 417 BY SIMILARITY.
FT DISULFID 419 428 BY SIMILARITY.
FT DISULFID 435 446 BY SIMILARITY.
FT DISULFID 440 455 BY SIMILARITY.
FT DISULFID 457 466 BY SIMILARITY.
SQ SEQUENCE 589 AA; 61424 MW; A17B3BF9B95EC17F CRC64;

Query Match 71.7%; Score 33; DB 1; Length 589;
Best Local Similarity 71.4%; Pred. No. 45;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 RVCVOPG 8
|||||

Db 59 RVCLKPG 65
ID DLL3_MOUSE STANDARD; PRT; 592 AA.
AC 088516; Q9QWZ7; Q35675; Q9QWZ9;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE DELTA-LIKE PROTEIN 3 PRECURSOR (DROSOPHILA DELTA HOMOLOG 3) (M-DELTA-
DE 3).
GN DLL3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6 X DBA; TISSUE=Embryo;
RX MEDLINE=97417575; PubMed=9272948;
RA Dunwoodie S.L., Henrique D.M.P., Harrison S.M., Beddington R.S.P.;
RT "Mouse Dll3: a novel divergent Delta gene which may complement the
RT function of other Delta homologues during early pattern formation in
RT the mouse embryo.";
RL Nat. Genet. 19:274-278(1998).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN=129/SVJ;
RX MEDLINE=98324780; PubMed=9662403;
RA Kusumi K., Sun E.S., Kerrebrock A.W., Bronson R.T., Chi D.-C.,
RA Bulotsky M.S., Spencer J.B., Birren B.W., Frankel W.N., Lander E.S.;
RT "The mouse pudgy mutation disrupts Delta homologue Dll3 and initiation
RT of early somite boundaries.";
RL Nat. Genet. 19:274-278(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Neural tube;
RA Nakayama K., Nakayama N., Tomooka Y., Hayashi Y., Takahashi M.;
RT "Specific expression of a divergent type of Delta in a set of earliest
RT generated neurons including the prospective subplate neurons.";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INHIBITS PRIMARY NEUROGENESIS. MAY BE REQUIRED TO DIVERT
CC NEURONS ALONG A SPECIFIC DIFFERENTIATION PATHWAY. PLAY A ROLE IN
CC THE FORMATION OF SOMITE BOUNDARIES DURING SEGMENTATION OF THE
CC PARAXIAL MESODERM.
CC -1- SUBUNIT: CAN BIND AND ACTIVATE NOTCH-1 OR ANOTHER NOTCH RECEPTOR
CC (PROBABLE).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PROBABLE).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 AND 2 (SHOWN HERE); ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE NEUROECTODERM
CC AND PARAXIAL MESODERM DURING EMBRYOGENESIS.
CC -1- DOMAIN: THE DELTA-SERRATE-LAG2 (DSL) DOMAIN IS REQUIRED FOR
CC BINDING TO THE NOTCH RECEPTOR.
CC -1- DISEASE: A TRUNCATING MUTATION IN DLL3 IS THE CAUSE OF THE PUDGY
CC (PU) PHENOTYPE. PUDGY MICE EXHIBIT PATTERNING DEFECTS AT THE
CC EARLIEST STAGES OF SOMITOGENESIS. ADULT PUDGY MICE PRESENT SEVERE
CC VERTEBRAL AND RIB DEFORMITIES.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF068865; AAC40170.1; -;
DR EMBL; AF068865; AAC40169.1; -;
DR EMBL; Y11895; CAA72637.1; -;

DR EMBL; AB013440; BAA33716.1; -
 DR HSP; P00740; 11XA.
 DR InterPro; IPR000561; -
 DR InterPro; IPR000742; -
 DR InterPro; IPR001438; -
 DR Pfam; PF00008; EGF; 6.
 DR PRINTS; PRO0010; EGFBL00D.
 DR PROSITE; PS00022; EGF_1; 6.
 DR PROSITE; PS01186; EGF_2; 6.
 DR DR; MGI:1096877; D113.
 KW Signal; EGF-like domain; Repeat; Transmembrane; Developmental protein;
 FT Differentiation; Alternative splicing.
 FT SIGNAL 1 32
 FT CHAIN 33 592
 FT DOMAIN 33 490
 FT TRANSMEM 491 511
 FT DOMAIN 512 592
 FT DOMAIN 174 213
 FT DOMAIN 235 246
 FT DOMAIN 296 307
 FT DOMAIN 337 348
 FT DOMAIN 375 386
 FT DOMAIN 413 424
 FT DOMAIN 451 462
 FT DISULFID 237 246
 FT DISULFID 281 296
 FT DISULFID 298 307
 FT DISULFID 325 337
 FT DISULFID 339 348
 FT DISULFID 355 366
 FT DISULFID 360 375
 FT DISULFID 377 386
 FT DISULFID 393 404
 FT DISULFID 398 413
 FT DISULFID 415 424
 FT DISULFID 431 442
 FT DISULFID 436 451
 FT DISULFID 453 462
 FT VARSPLIC 585 592
 FT CONFLICT 94 94
 FT CONFLICT 401 401
 FT SEQUENCE 592 AA; 62069 MW; 1A84F8022E7E7DCC CRC64;
 Query Match 71.7%; Score 33; DB 1; Length 592;
 Best Local Similarity 71.4%; Pred. No. 46;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RVCVQPG 8
 DB 59 RVCLKPG 65
 RESULT 13
 ID DPO2_HUMAN STANDARD; PRT; 598 AA.
 AC Q14181;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE DNA POLYMERASE ALPHA 70 KDA SUBUNIT (DNA POLYMERASE SUBUNIT B).
 GN POLA2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cervical epithelium;
 RX MEDLINE=94038939; PubMed=8223465;
 RA Collins K.L., Russo A.A.R., Tseng B.Y., Kelly T.J.;
 RT "The role of the 70 kDa subunit of human DNA polymerase alpha in DNA replication.";

RL EMBL J. 12:4555-4566(1993).
 RN [2]
 RP PHOSPHORYLATION.
 RX MEDLINE=91210318; PubMed=1902230;
 RA Nasheuer H.-P., Moore A., Wahl A.F., Wang T.S.-F.;
 RT "Cell cycle-dependent phosphorylation of human DNA polymerase alpha.";
 RL J. Biol. Chem. 266:7893-7903(1991).
 CC -!- FUNCTION: MAY PLAY AN ESSENTIAL ROLE AT THE EARLY STAGE OF
 CC CHROMOSOMAL DNA REPLICATION BY COUPLING THE POLYMERASE
 CC ALPHA/PRIMASE COMPLEX TO THE CELLULAR REPLICATION MACHINERY (BY
 CC SIMILARITY).
 CC -!- SUBUNIT: DNA POLYMERASE ALPHA-PRIMASE IS A FOUR SUBUNIT ENZYME
 CC (SUBUNITS A, B, C AND D), WHICH IS ASSEMBLED THROUGHOUT THE CELL
 CC CYCLE. THE LARGEST SUBUNIT (SUBUNIT A) HAS DNA POLYMERASE
 CC ACTIVITY, THE TWO SMALLEST SUBUNITS (SUBUNITS C AND D) HAVE DNA
 CC PRIMASE ACTIVITY. SUBUNIT B BINDS TO SUBUNIT A.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- DOMAIN: THE N-TERMINAL 240 AMINO ACIDS ARE SUFFICIENT TO MEDIATE
 CC COMPLEX FORMATION.
 CC -!- PTM: PHOSPHORYLATED IN A CELL CYCLE-DEPENDENT MANNER, IN G2/M
 CC PHASE.
 CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE ALPHA SUBUNIT B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL; L24559; AAA16459.1; -
 KW DNA replication; Nuclear protein; Phosphorylation.
 FT DOMAIN 101 107 POLY-GLU
 FT DOMAIN 115 157 PRO/SER/THR-RICH (HYDROPHILIC).
 FT DOMAIN 486 489 POLY-SER.
 SQ SEQUENCE 598 AA; 65977 MW; F2ED8D6BDF4751A3 CRC64;
 Query Match 71.7%; Score 33; DB 1; Length 598;
 Best Local Similarity 75.0%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GRVCVQPG 8
 DB 553 GCVCVNP 560
 RESULT 14
 ID DPO2_MOUSE STANDARD; PRT; 600 AA.
 AC P33611;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE DNA POLYMERASE ALPHA 70 KDA SUBUNIT (DNA POLYMERASE SUBUNIT B).
 GN POLA2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 84-102; 269-285 AND 394-403.
 RX MEDLINE=93216788; PubMed=8463324;
 RA Miyazawa H., Izumi M., Tada S., Takada R., Masutani M., U1 M.,
 RA Hanaoka F.;
 RT "Molecular cloning of the cDNAs for the four subunits of mouse DNA
 RT polymerase alpha-primase complex and their gene expression during
 RT cell proliferation and the cell cycle.";
 RL J. Biol. Chem. 268:8111-8122(1993).
 CC -!- FUNCTION: MAY PLAY AN ESSENTIAL ROLE AT THE EARLY STAGE OF
 CC CHROMOSOMAL DNA REPLICATION BY COUPLING THE POLYMERASE
 CC ALPHA/PRIMASE COMPLEX TO THE CELLULAR REPLICATION MACHINERY (BY

CC CC SIMILARITY).

CC CC (-1- SUBUNIT: DNA POLYMERASE ALPHA-PRIMASE IS A FOUR SUBUNIT ENZYME

CC CC (SUBUNITS A, B, C AND D), WHICH IS ASSEMBLED THROUGHOUT THE CELL

CC CC CYCLE. THE LARGEST SUBUNIT (SUBUNIT A) HAS DNA POLYMERASE

CC CC ACTIVITY, THE TWO SMALLEST SUBUNITS (SUBUNITS C AND D) HAVE DNA

CC CC PRIMASE ACTIVITY. SUBUNIT B BINDS TO SUBUNIT A.

CC CC (-1- SUBCELLULAR LOCATION: NUCLEAR.

CC CC (-1- PTM: PHOSPHORYLATED IN A CELL CYCLE-DEPENDENT MANNER, IN G2/M

CC CC PHASE (BY SIMILARITY).

CC CC (-1- SIMILARITY: BELONGS TO THE DNA POLYMERASE ALPHA SUBUNIT B FAMILY.

CC CC -----

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC CC the European Bioinformatics Institute. There are no restrictions on its

CC CC use by non-profit institutions as long as its content is in no way

CC CC modified and this statement is not removed. Usage by and for commercial

CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC CC or send an email to license@isb-sib.ch).

CC CC -----

CC CC EMBL; D13546; BAA02746.1; -

CC CC PIR; B46642; B46642.

CC CC MGI; 99690; Pola2.

CC CC DNA replication; Nuclear protein; Phosphorylation.

CC CC FT DOMAIN 101 107 POLY-GLU.

CC CC FT DOMAIN 115 137 PRO/SER/THR-RICH (HYDROPHILIC).

CC CC SEQUENCE 600 AA; 66267 MW; 79F94BE6EF33FBC CRC64;

CC CC -----

CC CC Query Match 71.7%; Score 33; DB 1; Length 600;

CC CC Best Local Similarity 75.0%; Pred. No. 45;

CC CC Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8

DB 553 GCVCVNP 560

RESULT 15

DL33_HUMAN

ID DL33_HUMAN STANDARD; PRT; 618 AA.

AC Q9NVJ7;

DT 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE DELTA-LIKE PROTEIN 3 PRECURSOR (DROSOPHILA DELTA HOMOLOG 3).

GN DL33.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND VARIANT SD ASP-385.

RX MEDLINE=20206573; PubMed=10742114;

RA Bulman M.P., Kusumi K., Frayling T.M., McKeown C., Garrett C.,

RA Lander E.S., Krumlauf R., Hattersley A.T., Ellard S., Turpenny P.D.;

RT "Mutations in the human delta homologue, DL33, cause axial skeletal

RT defects in spondylocostal dysostosis."

RL Nat. Genet. 24:438-441(2000).

CC CC (-1- FUNCTION: INHIBITS PRIMARY NEUROGENESIS. MAY BE REQUIRED TO DIVERT

CC CC NEURONS ALONG A SPECIFIC DIFFERENTIATION PATHWAY. PLAY A ROLE IN

CC CC THE FORMATION OF SOMITE BOUNDARIES DURING SEGMENTATION OF THE

CC CC PARAXIAL MESODERM (BY SIMILARITY).

CC CC (-1- SUBUNIT: CAN BIND AND ACTIVATE NOTCH-1 OR ANOTHER NOTCH RECEPTOR

CC CC (BY SIMILARITY).

CC CC (-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PROBABLE).

CC CC (-1- DOMAIN: THE DELTA-SERRATE-LAG2 (DSL) DOMAIN IS REQUIRED FOR

CC CC BINDING TO THE NOTCH RECEPTOR.

CC CC (-1- DISEASE: DEFECTS IN DL33 ARE A CAUSE OF AUTOSOMAL RECESSIVE

CC CC SPONDYLOCOSTAL DYSOSTOSIS (SD). IT IS CHARACTERIZED BY MULTIPLE

CC CC HEMIVERTEBRAE, RIB FUSIONS AND DELETIONS WITH A NON-PROGRESSIVE

CC CC KYPHOSCOLIOSIS.

CC CC (-1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.

CC CC (-1- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC CC the European Bioinformatics Institute. There are no restrictions on its

CC CC use by non-profit institutions as long as its content is in no way

CC CC modified and this statement is not removed. Usage by and for commercial

CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC CC or send an email to license@isb-sib.ch).

CC CC -----

CC CC EMBL; AF241373; AAF62542.1; -

CC CC EMBL; AF241367; AAF62542.1; JOINED.

CC CC EMBL; AF241368; AAF62542.1; JOINED.

CC CC EMBL; AF241369; AAF62542.1; JOINED.

CC CC EMBL; AF241370; AAF62542.1; JOINED.

CC CC EMBL; AF241371; AAF62542.1; JOINED.

CC CC EMBL; AF241372; AAF62542.1; JOINED.

CC CC MIM; 602768; -

CC CC MIM; 277300; -

CC CC InterPro; IPR000561; -

CC CC InterPro; IPR000742; -

CC CC InterPro; IPR001438; -

CC CC PROSITE; PS00022; EGF_1; 6.

CC CC PROSITE; PS01186; EGF_2; 6.

CC CC Signal; EGF-like domain; Repeat; Transmembrane; Developmental protein;

CC CC Differentiation; Disease mutation.

CC CC FT SIGNAL 1 26

CC CC FT CHAIN 27 618

CC CC FT DOMAIN 27 492

CC CC FT DOMAIN 493 513

CC CC FT DOMAIN 514 618

CC CC FT DOMAIN 176 215

CC CC FT DOMAIN 237 248

CC CC FT DOMAIN 298 309

CC CC FT DOMAIN 339 350

CC CC FT DOMAIN 377 388

CC CC FT DOMAIN 415 426

CC CC FT DOMAIN 453 464

CC CC FT DISULFID 239 248

CC CC FT DISULFID 283 298

CC CC FT DISULFID 309 309

CC CC FT DISULFID 327 339

CC CC FT DISULFID 341 350

CC CC FT DISULFID 357 368

CC CC FT DISULFID 362 377

CC CC FT DISULFID 379 388

CC CC FT DISULFID 395 406

CC CC FT DISULFID 400 415

CC CC FT DISULFID 417 426

CC CC FT DISULFID 433 444

CC CC FT DISULFID 438 453

CC CC FT DISULFID 455 464

CC CC FT VARIANT 385

CC CC SEQUENCE 618 AA; 64617 MW; 58A9BC0A7DEAD1A0 CRC64;

CC CC /FTid=VAR_009952.

Query Match 71.7%; Score 33; DB 1; Length 618;

Best Local Similarity 71.4%; Pred. No. 47;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVCVQPG 8

DB 61 RVCLKPG 67

Search completed: June 13, 2001, 14:21:39

Job time: 799 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: June 13, 2001, 14:08:20 ; Search time 225.85 Seconds
(without alignments)
4.152 Million cell updates/sec

Title: PCT-US01-05825A-1
Perfect score: 46
Sequence: 1 GRVCVQPG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mmc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_unclassified:*

13: sp_vertebrate:*

14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	84.8	4123	4 075851	075851 homo sapien
2	38	82.6	783	14 068952	068952 hepatitis c
3	37	80.4	447	2 086673	086673 streptomyce
4	37	80.4	1894	11 064487	064487 mus musculus
5	35	76.1	236	10 09XHV8	09XHV8 oryza sativ
6	35	76.1	323	2 09L8F5	09L8F5 vibrio mimi
7	35	76.1	399	2 09R3V6	09R3V6 vibrio chol
8	35	76.1	399	2 09L7Q6	09L7Q6 vibrio chol
9	35	76.1	578	5 09W095	09W095 drosophila
10	35	76.1	680	11 09QW15	09QW15 mus sp. bet
11	35	76.1	707	2 077907	077907 moorella th
12	35	76.1	723	11 09QW16	09QW16 rattus sp.
13	34	73.9	96	2 087588	087588 rhizobium l
14	34	73.9	217	10 040295	040295 pycnococcus
15	34	73.9	338	2 059545	059545 morganella
16	34	73.9	371	1 09UX13	09UX13 sulfolobus
17	34	73.9	465	8 032591	032591 mazzella l
18	34	73.9	593	14 087093	087093 simian herp
19	34	73.9	784	4 015495	015495 homo sapien

20	34	73.9	784	6	097702	097702 canis famil
21	34	73.9	784	6	09TUN7	09TUN7 canis famil
22	34	73.9	784	6	09TUNS	09TUNS sus scrofa
23	34	73.9	788	6	09TUN3	09TUN3 oryctolagus
24	34	73.9	1827	5	020535	020535 caenorhabdi
25	33	71.7	139	11	088515	088515 mus musculu
26	33	71.7	213	5	044899	044899 caenorhabdi
27	33	71.7	226	5	045327	045327 caenorhabdi
28	33	71.7	265	5	09W452	09W452 drosophila
29	33	71.7	273	4	09UHFI	09UHFI homo sapien
30	33	71.7	294	5	09VV62	09VV62 drosophila
31	33	71.7	298	10	09T091	09T091 arabidopsis
32	33	71.7	416	10	09M8J4	09M8J4 arabidopsis
33	33	71.7	440	5	044637	044637 caenorhabdi
34	33	71.7	454	4	014060	014060 homo sapien
35	33	71.7	497	11	09WUJ0	09WUJ0 mus musculu
36	33	71.7	499	6	09TSQ9	09TSQ9 cercopithec
37	33	71.7	585	11	035675	035675 mus musculu
38	33	71.7	585	11	09QWZ7	09QWZ7 mus musculu
39	33	71.7	585	11	09QWL9	09QWL9 mus musculu
40	33	71.7	589	11	088671	088671 rattus norv
41	33	71.7	590	11	09JJPI	09JJPI mus musculu
42	33	71.7	592	11	088516	088516 mus musculu
43	33	71.7	611	2	069514	069514 mycobacteri
44	33	71.7	618	4	09NYJ7	09NYJ7 homo sapien
45	33	71.7	631	11	09JJPI2	09JJPI2 mus musculu

ALIGNMENTS

RESULT 1
075851
ID 075851 PRELIMINARY; PRT; 4123 AA.
AC 075851;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE WUGSC:H-DJ0751H13.1 PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Leonard S., Graves T., Stromatt C.;
RT "The sequence of Homo sapiens PAC clone DJ0751H13.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Waterston R.H.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004877; AAC36301.1; -
DR HSP; P01130; 1AJJ
DR INTERPRO; IPR000421; -
DR INTERPRO; IPR000561; -
DR INTERPRO; IPR000884; -
DR INTERPRO; IPR000923; -
DR INTERPRO; IPR001007; -
DR INTERPRO; IPR001064; -
DR INTERPRO; IPR001092; -
DR INTERPRO; IPR001846; -
DR INTERPRO; IPR002172; -
DR INTERPRO; IPR002223; -
DR INTERPRO; IPR002465; -
DR INTERPRO; IPR002919; -
DR PFAM; PF00057; 1dl_recept_a; 11.
DR PFAM; PF00090; tsp_1; 14.
DR PFAM; PF00094; vwd; 5.

DR PFAM; PF00754; F5_F8_type_C; 1.
 DR PFAM; PF01826; TIL; 9.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS00196; COPPER_BLUE; UNKNOWN_1.
 DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
 DR PROSITE; PS00280; BPTI_KUNITZ; UNKNOWN_1.
 DR PROSITE; PS00340; RECEPTOR_CYTOKINES_2; UNKNOWN_6.
 DR PROSITE; PS01209; LDLRA_1; 9.
 DR PROSITE; PS00068; LDLRA_2; 20.
 KW Glycoprotein.
 FT NON_TER 1
 SQ SEQUENCE 4123 AA; 434985 MW; 7AAB6F8DCE012FB CRC64;

Query Match 84.8%; Score 39; DB 4; Length 4123;
 Best Local Similarity 75.0%; Pred. No. 71;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVOPG 8
 | : |||||
 Db 3234 GAICVOPG 3241

RESULT 2

ID Q68952 PRELIMINARY; PRT; 783 AA.
 AC Q68952;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE (JK4) CORE, EL, NS1/E2 AND NS2 GENES (FRAGMENTS).
 GN JK4.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PATIENT 4:HEPATITIS C VIRUS, JK4;
 RX MEDLINE=93119270; PubMed=8380322;
 RA Honda M., Kaneko S., Unoura M., Kobayashi K., Murakami S.:
 RT "Sequence analysis of putative structural regions of hepatitis C virus
 RT isolated from 5 Japanese patients with hepatocellular carcinoma.";
 RL Arch. Virol. 128:163-169(1993).
 DR EMBL; X61594; CAA43791.1; -. CORE.
 FT CHAIN 1 >191
 FT NON_CONS 191 192
 FT CHAIN 192 >383
 FT NON_CONS 383 384
 FT CHAIN 384 >733
 FT NON_CONS 733 734
 FT CHAIN 734 >783
 FT NON_TER 783 783
 SQ SEQUENCE 783 AA; 85808 MW; A3145A3D310F9E5C CRC64;

Query Match 82.6%; Score 38; DB 14; Length 783;
 Best Local Similarity 75.0%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVOPG 8
 || | |||
 Db 73 GRACVOPG 80

RESULT 3

ID Q86673 PRELIMINARY; PRT; 447 AA.
 AC Q86673;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)

DE PUTATIVE LIPOPROTEIN.
 GN SC4A2.17C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Oliver K., Harris D.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Parkhill J., Barrell B.G., Raftery M.A.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RX MEDLINE=9700351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL031182; CAA20169.1; -.
 KW Lipoprotein.
 SQ SEQUENCE 447 AA; 46712 MW; 809E0091B7834D80 CRC64;

Query Match 80.4%; Score 37; DB 2; Length 447;
 Best Local Similarity 85.7%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVOP 7
 ||:||||
 Db 31 GRUCVOP 37

RESULT 4

ID Q64487 PRELIMINARY; PRT; 1894 AA.
 AC Q64487; Q64486; Q64488; Q64495;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE PROTEIN-TYROSINE PHOSPHATASE, RECEPTOR-TYPE, D PRECURSOR (EC 3.1.3.48)
 DE (PROTEIN-TYROSINE PHOSPHATASE DELTA) (R-PTP-DELTA).
 DE PTPRD.
 GN Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-BRAIN;
 RX MEDLINE=93360986; PubMed=8355697;
 RA Mizuno K., Hasegawa K., Katagiri T., Ogimoto M., Ichikawa T.,
 RA Yakura H.;
 RT "Mpp delta, a putative murine homolog of HTPP delta, is expressed in
 RT specialized regions of the brain and in the B-cell lineage.";
 RL Mol. Cell. Biol. 13:5513-5523(1993).
 RN [2]
 RP SEQUENCE OF 1430-1534 FROM N.A.
 RC STRAIN-BALB/C; TISSUE-BRAIN;
 RX MEDLINE=95134232; PubMed=7832766;
 RA Hendriks W., Schepens J., Brugman C., Zeeuwen P., Wieringa B.;
 RT "A novel receptor-type protein tyrosine phosphatase with a single
 RT catalytic domain is specifically expressed in mouse brain.";
 RL Biochem. J. 305:499-504(1995).
 CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O = PROTEIN
 CC TYROSINE + PHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -1- ALTERNATIVE PRODUCTS: AT LEAST THREE DIFFERENT ISOFORMS; TYPE A,
CC B, AND C (SHOWN HERE); ARE DUE TO ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: BRAIN, KIDNEY, HEART, AND SOME B-CELL LINES.
CC -1- PTM: A CLEAVAGE OCCURS THAT SEPARATES THE EXTRACELLULAR DOMAIN
CC FROM THE TRANSMEMBRANE SEGMENT (BY SIMILARITY).
CC -1- SIMILARITY: EXTRACELLULAR REGION TYPICAL OF A CAM FAMILY (3 IG-
CC LIKE DOMAINS FOLLOWED BY 8 FIBRONECTIN TYPE III-LIKE DOMAINS), AND
CC A CYTOPLASMIC REGION COMPOSED OF TWO PTPASE DOMAINS.
DR EMBL; D13903; BAA03003.1; -;
DR EMBL; D13905; BAA03005.1; -;
DR EMBL; D13904; BAA03004.1; -;
DR EMBL; D23051; CAA05056.1; -;
DR HSPF; P18052; 11FO.
DR MGD; MGI:97812; Ptpcd.
DR INTERPRO; IPR000242; -;
DR INTERPRO; IPR000387; -;
DR INTERPRO; IPR001777; -;
DR INTERPRO; IPR003006; -;
DR PFAM; PF00041; fn3; 3;
DR PFAM; PF00047; ig; 1;
DR PFAM; PF00102; Y_phosphatase; 2;
DR PRINTS; PR00014; ENTPEPIII.
DR PRINTS; PR00700; PTPPHPTASE.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 4;
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 4;
KW Hydrolase; Receptor; Glycoprotein; Signal; Transmembrane; Duplication;
KW Immunoglobulin domain; Alternative splicing; Repeat.
FT SIGNAL 1 18
FT CHAIN 19 1894
FT DOMAIN 19 1251
FT TRANSMEM 1252 1272
FT DOMAIN 1273 1894
FT DOMAIN 45 112
FT DOMAIN 147 215
FT DOMAIN 245 304
FT DOMAIN 315 409
FT DOMAIN 412 508
FT DOMAIN 511 601
FT DOMAIN 603 702
FT DOMAIN 705 806
FT DOMAIN 809 899
FT DOMAIN 901 1000
FT DOMAIN 1004 1122
FT DOMAIN 1360 1601
FT DOMAIN 1602 1894
FT DISULFID 52 105
FT DISULFID 154 208
FT DISULFID 252 297
FT ACT_SITE 1536 1536
FT ACT_SITE 1826 1826
FT SITE 1160 1163
FT VARSPPLIC 1 1
FT VARSPPLIC 2 225
FT VARSPPLIC 602 1001
FT CONFLICT 1430 1432
FT CONFLICT 1475 1485
FT CONFLICT 1493 1493
FT CONFLICT 1514 1516
FT CONFLICT 1519 1525
SQ SEQUENCE 1894 AA; 212194 MW; 8211033B53DD163C CRC64;

Query Match 80.4%; Score 37; DB 11; Length 1894;
Best Local Similarity 85.7%; Pred. No. 85;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GRVCVQP 7
Db 1667 GRVCLOP 1673

RESULT 5
O9XHV8 PRELIMINARY; PRT; 236 AA.
AC O9XHV8;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE 10A191.8.
GN 10A191.8.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LEMONT;
RA Vysotskaya V.S., Schwartz J.R., Osborne B.I., Wing R., Yu G., Kwan A.,
RA Liu S., Lee J., Toriumi M., Luros J., Li J., Kremenetskaia I., Oji O.,
RA Theologis A.;
RT "Oryza sativa chromosome 1 BAC 10A191.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007858; AAD39593.1; -;
DR INTERPRO; IPR001005; -;
SQ SEQUENCE 236 AA; 26681 MW; 0A33BE5977933E42 CRC64;

Query Match 76.1%; Score 35; DB 10; Length 236;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 VCVQPG 8
Db 227 VCVQPG 232

RESULT 6
O9L8F5 PRELIMINARY; PRT; 323 AA.
AC O9L8F5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE ZOT (FRAGMENT).
GN ZOT.
OS Vibrio mimicus.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=674;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PT5;
RX MEDLINE=20143766; PubMed=10678967;
RA Boyd E.F., Moyer K.E., Shi L., Waldo M.K.;
RT "Infectious CTXPhi and the vibrio pathogenicity island prophage in
RT vibrio mimicus: evidence for recent horizontal transfer between V.
RT mimicus and V. cholerae.";
RL Infect. Immun. 68:1507-1513(2000).
DR EMBL; AF207857; AAF40142.1; -;
FT NON_TER 1 1
FT NON_TER 323 323
SQ SEQUENCE 323 AA; 36306 MW; 01C12DAE9B873C3B CRC64;

Query Match 76.1%; Score 35; DB 2; Length 323;
Best Local Similarity 75.0%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GRVCVQP 8
Db 240 GRVCVQD 247

RESULT 7

Q9R3V6
ID Q9R3V6 PRELIMINARY; PRT; 399 AA.
AC Q9R3V6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ZONULAR OCCLUDENS TOXIN (ZONA OCCLUDENS TOXIN).
GN ZOT OR VC1458.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KNIH002;
RA Shin H.J., Park Y.C., Kim Y.C.;
RT "Cloning and nucleotide sequence analysis of the virulence gene
RT cassette from Vibrio cholerae KNIH002 isolated in Korea.";
RT Misalunurhag Hoiji 35:205-210(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=0139-TOR OGAWA;
RA zhi-Yong H., Wei-Jie Z., Xiang-Fu W.;
RT "Cloning and Expression of zot Gene from Vibrio cholerae.";
RT Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heideberg J.F., Elsen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Hart D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RT Nature 406:477-483(2000).
RL EMBL; AF175708; AAD51358.1; -
DR EMBL; AF123049; AAD26854.1; -
DR EMBL; AE004224; AAF94615.1; -
DR TIGR; VC1458; -
SQ SEQUENCE 399 AA; 44903 MW; 3C7424B758176774 CRC64;

Query Match 76.1%; Score 35; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8
DB 291 GRLCVQDG 298
||:|||||

RESULT 8
Q9L7Q6 PRELIMINARY; PRT; 399 AA.
ID Q9L7Q6;
AC Q9L7Q6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ZOT.
GN ZOT.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=86015;
RA Kan B., Liu Y.Q., Qi G.M., Gao S.Y.;
RT "Vibrio cholerae nct-crxphi whole genome, include rstr(Rstr),
RT rsta(Rsta), rstB(RstB), cep(Cep), orfU(OrfU), ace(Ace) and zot(Zot)
RT genes.";

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF220606; AAF29547.1; -
SQ SEQUENCE 399 AA; 44990 MW; CFA3DBCC9E23EE1 CRC64;

Query Match 76.1%; Score 35; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8
DB 291 GRLCVQDG 298
||:|||||

RESULT 9
Q9W095 PRELIMINARY; PRT; 578 AA.
ID Q9W095;
AC Q9W095; Q9W094;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE CG7995 PROTEIN.
GN CG7995;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Nusskern D.R., Pacht J.M.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacht J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (AC Q9VP61) AND
CC SHORT ISOFORM (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
DR EMBL; AE003472; AAF47558.1; -


```

DR EMBL; AE003472; AAF47559.1; -.
DR HSSP; P08859; 1GLC.
DR FLYBASE; FBgn0035266; CG7995.
DR INTERPRO; IPR000577; -.
DR PFAM; PF00370; FGCY; 1.
KW Alternative splicing; Hypothetical protein.
FT VARSPPLIC 1 2 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 578 AA; 64905 MW; D77429D03B65D8D3 CRC64;

Query Match 76.1%; Score 35; DB 5; Length 578;
Best Local Similarity 62.5%; Pred. No. 71;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8
Db 286 GQMCVKRPG 293
:::|:|:|

RESULT 10
Q9QW15 ID Q9QW15 PRELIMINARY; PRT; 680 AA.
AC Q9QW15;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE BETA 3 INTEGRIN, GPIIIB.
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=93290675; PubMed=8512576;
RA Cleut A.M., Rosa J.P., Letourneur F., Poncz M., Rifat S.;
RT "A comparative analysis of cDNA-derived sequences for rat and mouse
beta 3 integrins (GPIIIB) with their human counterpart.";
RL Biochem. Biophys. Res. Commun. 193:771-778(1993).
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR001169; -.
DR INTERPRO; IPR002086; -.
DR INTERPRO; IPR002369; -.
DR PFAM; PF00362; Integrin_B; 1.
DR PRINTS; PR01186; INTEGRINB.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00243; INTEGRIN_BETA; 3.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
SQ SEQUENCE 680 AA; 75514 MW; 672DB7338DD86003 CRC64;

Query Match 76.1%; Score 35; DB 11; Length 680;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCQVQPG 8
Db 509 VCQVQPG 514
|||||

RESULT 11
P77907 ID P77907 PRELIMINARY; PRT; 707 AA.
AC P77907;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE FORMATE DEHYDROGENASE, BETA SUBUNIT (EC 1.2.1.43) (FORMATE
DE DEHYDROGENASE [NADP+]).
GN FDHB.
OS Moorella thermoacetica (Clostridium thermoacetum).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Moorella group;
OC Moorella.

```

```

OX NCBI_TaxID=1525;
RN [1]
SEQUENCE FROM N.A.
RA Li X.L., Ljungdahl L.G., Gollin D.J.;
RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDS NADP.
CC -1- CATALYTIC ACTIVITY: FORMATE + NADP(+) -> CO(2) + NADPH.
CC -1- COFACTOR: SELENIUM; TUNGSTEN; IRON.
DR EMBL; U73807; AAB18329.1; -.
DR INTERPRO; IPR000205; -.
DR INTERPRO; IPR001100; -.
DR INTERPRO; IPR001949; -.
DR PFAM; PF00070; pyr_redox; 1.
DR PROSITE; PS00645; COMPLEX1_51K_2; 1.
KW Oxidoreductase; Transmembrane; NADP.
FT TRANSMEM 38 58 POTENTIAL.
FT DOMAIN 178 181 POLY-ALA.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 540 560 POTENTIAL.
FT TRANSMEM 576 596 POTENTIAL.
SQ SEQUENCE 707 AA; 75057 MW; 3598D2D0CCF6797E CRC64;

Query Match 76.1%; Score 35; DB 2; Length 707;
Best Local Similarity 85.7%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVQPG 7
Db 237 GRVCVHP 243
|||||

RESULT 12
Q9QW16 ID Q9QW16 PRELIMINARY; PRT; 723 AA.
AC Q9QW16;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE BETA 3 INTEGRIN, GPIIIB.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=93290675; PubMed=8512576;
RA Cleut A.M., Rosa J.P., Letourneur F., Poncz M., Rifat S.;
RT "A comparative analysis of cDNA-derived sequences for rat and mouse
beta 3 integrins (GPIIIB) with their human counterpart.";
RL Biochem. Biophys. Res. Commun. 193:771-778(1993).
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR001169; -.
DR INTERPRO; IPR002086; -.
DR INTERPRO; IPR002369; -.
DR PFAM; PF00362; Integrin_B; 1.
DR PRINTS; PR01186; INTEGRINB.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00243; INTEGRIN_BETA; 3.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
SQ SEQUENCE 723 AA; 80233 MW; B31D3127E6ECF90F CRC64;

Query Match 76.1%; Score 35; DB 11; Length 723;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCQVQPG 8
Db 548 VCQVQPG 553
|||||

```

Db 142 GMYCQVP 148

RESULT 15

Q59545

ID Q59545 PRELIMINARY; PRT; 338 AA.

AC Q59545

DT 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)

DE XYLITOL DEHYDROGENASE (EC 1.1.1.9) (D-XILOULOSE REDUCTASE).

OS Morganella morganii (Proteus morganii).

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Morganella.

OC NCBI_TaxID=582;

OX [1]

RN SEQUENCE FROM N.A.

RP Gallo M.A.;

RL Thesis (1991), University of Wisconsin-Madison, Madison, USA.

CC -1- CATALYTIC ACTIVITY: XYLITOL + NAD(+) -> D-XILOULOSE + NADH.

DR EMBL: L34345; AAA25324.1; -

DR HSSP: P07846; 1SDG

DR INTERPRO: IPR002085; -

DR INTERPRO: IPR002328; -

DR PFAM: PF00107; adh_zinc; 1.

DR PROSITE: PS00059; ADH_ZINC; 1.

KW Oxidoreductase.

FT NON_TER 96

SQ SEQUENCE 338 AA; 35951 MW; 4C6119553DC51873 CRC64;

Query Match 73.9%; Score 34; DB 2; Length 338;

Best Local Similarity 71.4%; Pred. No. 69;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVCVQPG 8

DB 86 RVCMEPG 92

Search completed: June 13, 2001, 14:20:21

Job time: 721 sec

RESULT 13

O87588

ID O87588 PRELIMINARY; PRT; 96 AA.

AC O87588

DT 01-NOV-1998 (Tremblrel. 08, Created)

DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)

DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)

DE SORBITOL DEHYDROGENASE (FRAGMENT).

GN SRLD.

OS Rhizobium leguminosarum (biovar trifolii).

OG Plasmid pRleW14-2c.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Rhizobium.

OC NCBI_TaxID=386;

OX [1]

RN SEQUENCE FROM N.A.

RP STRAIN=RLT100;

RA Oresnik I.J., Pacarynk L.A., O'Brien S.A.P., Yost C.K., Hynes M.F.;

RT "Plasmid Encoded Catabolic Genes in Rhizobium leguminosarum bv. trifolii: Evidence for a plant inducible rhamnose locus involved in competition for nodulation."

RL Mol. Plant Microbe Interact. 0:0-0(1998).

DR EMBL: AF086782; AAD11983.1; -

DR HSSP: P07846; 1SDG

DR INTERPRO: IPR002328; -

DR PROSITE: PS00059; ADH_ZINC; 1.

KW Plasmid; Oxidoreductase; Zinc.

FT NON_TER 96

SQ SEQUENCE 96 AA; 10288 MW; D809EC6E706DFF43 CRC64;

Query Match 73.9%; Score 34; DB 2; Length 96;

Best Local Similarity 71.4%; Pred. No. 22;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVCVQPG 8

DB 87 RVCMEPG 93

RESULT 14

Q40295

ID Q40295 PRELIMINARY; PRT; 217 AA.

AC Q40295

DT 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)

DE T3/T7-LIKE RNA POLYMERASE (FRAGMENT).

GN RPOT.

OS Pycnococcus provasolii.

OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae; Mamiellales;

OC Pycnococaceae; Pycnococcus.

OX NCBI_TaxID=41880;

RN [1]

RN SEQUENCE FROM N.A.

RP STRAIN=GUILLARD (1991) CCMP 1203;

RX MEDLINE=96177564; PubMed=8604305;

RA Cernakian N., Ikeda T.M., Cedergren R., Gray M.W.;

RT "Sequences homologous to yeast mitochondrial and bacteriophage T3 and T7 RNA polymerases are widespread throughout the eukaryotic lineage."

RL Nucleic Acids Res. 24:648-654(1996).

DR EMBL: U34286; AAB01080.1; -

DR MENDEL; 12251; Pycpr:Rpot;12251.

FT NON_TER 1

FT NON_TER 217

FT NON_TER 217

SQ SEQUENCE 217 AA; 24820 MW; 933F6D489F7B2FCD CRC64;

Query Match 73.9%; Score 34; DB 10; Length 217;

Best Local Similarity 85.7%; Pred. No. 46;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVQVP 7

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:08:19 ; Search time 118.55 seconds
(without alignments)
1.296 Million cell updates/sec

Title: PCT-US01-05825A-1
Perfect score: 46
Sequence: 1 GRVCVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	76.1	400	2	US-08-624-601-8
2	34	73.9	92	2	US-07-728-215-39
3	34	73.9	92	2	US-07-728-215-41
4	34	73.9	718	1	US-08-444-792-4
5	34	73.9	718	1	US-08-445-042-4
6	34	73.9	788	2	US-07-728-215-32
7	32	69.6	76	3	US-08-851-843A-205
8	32	69.6	76	4	US-08-974-549A-324
9	32	69.6	118	2	US-08-652-816A-12
10	32	69.6	770	1	US-08-445-135-2
11	31	67.4	89	2	US-08-465-380-48
12	31	67.4	89	2	US-08-486-397-48
13	31	67.4	89	2	US-08-486-399-48
14	31	67.4	89	2	US-08-461-965-48
15	31	67.4	89	2	US-08-634-641-48
16	31	67.4	89	3	US-09-249-471-48
17	31	67.4	89	3	US-09-249-472-48
18	31	67.4	89	3	US-09-249-451-48
19	31	67.4	89	3	US-08-809-455-48
20	31	67.4	89	3	US-09-249-461-48
21	31	67.4	89	4	US-09-249-448-48
22	31	67.4	108	2	US-08-465-380-24
23	31	67.4	108	2	US-08-480-478-53
24	31	67.4	108	2	US-08-486-397-24
25	31	67.4	108	2	US-08-486-399-24
26	31	67.4	108	2	US-08-461-965-24
27	31	67.4	108	2	US-08-326-110A-53

28	31	67.4	108	2	US-08-634-641-24	Sequence 24, Appl
29	31	67.4	108	3	US-09-249-471-24	Sequence 24, Appl
30	31	67.4	108	3	US-09-249-472-24	Sequence 24, Appl
31	31	67.4	108	3	US-09-249-451-24	Sequence 24, Appl
32	31	67.4	108	3	US-08-809-455-24	Sequence 24, Appl
33	31	67.4	108	3	US-09-249-461-24	Sequence 24, Appl
34	31	67.4	108	4	US-09-249-448-24	Sequence 24, Appl
35	31	67.4	130	4	US-09-347-819-8	Sequence 8, Appl
36	31	67.4	171	2	US-08-465-380-62	Sequence 62, Appl
37	31	67.4	171	2	US-08-486-397-62	Sequence 62, Appl
38	31	67.4	171	2	US-08-486-399-62	Sequence 62, Appl
39	31	67.4	171	2	US-08-461-965-62	Sequence 62, Appl
40	31	67.4	171	2	US-08-634-641-62	Sequence 62, Appl
41	31	67.4	171	3	US-09-249-471-62	Sequence 62, Appl
42	31	67.4	171	3	US-09-249-472-62	Sequence 62, Appl
43	31	67.4	171	3	US-09-249-451-62	Sequence 62, Appl
44	31	67.4	171	3	US-08-809-455-62	Sequence 62, Appl
45	31	67.4	171	4	US-09-249-448-62	Sequence 62, Appl

ALIGNMENTS

RESULT 1
US-08-624-601-8
; Sequence 8, Application US/08624601
; Patent No. 5882653
; GENERAL INFORMATION:
; APPLICANT: Kaper Dr., James B.
; APPLICANT: Levine Dr., Myron M.
; TITLE OF INVENTION: Vibrio cholerae O1 (CVD111) and non-O1
; TITLE OF INVENTION: (CVD112 and CVD112RM) serogroup vaccine strains, methods
; TITLE OF INVENTION: of making same and products thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spencer & Frank
; STREET: 1100 New York Ave. N.W. Suite 300 East
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/624,601
; APPLICATION NUMBER: US/08/624,601
; FILING DATE: 08-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schneller Dr., John W.
; REGISTRATION NUMBER: 26,031
; REFERENCE/DOCKET NUMBER: BAWC20019P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)414-4000
; TELEFAX: (202)414-4040
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 400 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Vibrio cholerae
; STRAIN: El Tor 7946
; IMMEDIATE SOURCE:
; CLONE: zot
US-08-624-601-8

Query Match 76.1%; Score 35; DB 2; Length 400;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8
||:||||
Db 291 GRICVQDG 298

RESULT 2

US-07-728-215-39
; Sequence 39, Application US/07728215
; Patent No. 5962643
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Dean
; APPLICANT: Quaranta, Vito
; APPLICANT: Pytela, Robert
; TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States of America
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,215
; FILING DATE: 19910711
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P31 8717
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-728-215-39

Query Match 76.1%; Score 35; DB 2; Length 400;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8

||:||||

Db 291 GRICVQDG 298

RESULT 2

US-07-728-215-39
; Sequence 39, Application US/07728215
; Patent No. 5962643
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Dean
; APPLICANT: Quaranta, Vito
; APPLICANT: Pytela, Robert
; TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States of America
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,215
; FILING DATE: 19910711
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P31 8717
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-728-215-39

Query Match 76.1%; Score 35; DB 2; Length 400;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VCVOQPG 8
||:||||
Db 82 VCIQPG 87

RESULT 3

US-07-728-215-41
; Sequence 41, Application US/07728215
; Patent No. 5962643
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Dean
; APPLICANT: Quaranta, Vito
; APPLICANT: Pytela, Robert
; TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States of America
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,215
; FILING DATE: 19910711
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P31 8717
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-728-215-41

Query Match 73.9%; Score 34; DB 2; Length 92;

Best Local Similarity 83.3%; Pred. No. 20;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCVOQPG 8

||:||||

Db 82 VCIQPG 87

RESULT 4

US-08-444-792-4
; Sequence 4, Application US/08444792
; Patent No. 5726037
; GENERAL INFORMATION:
; APPLICANT: Bodary, Sarah C.
; APPLICANT: Gorman, John W.
; APPLICANT: McLean, John W.
; APPLICANT: Napier, Mary A.
; TITLE OF INVENTION: METHOD FOR PREPARING WATER SOLUBLE
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,792
; FILING DATE: 19-May-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380227
; FILING DATE: 30-JAN-1995
; PRIOR APPLICATION DATA:

Query Match 73.9%; Score 34; DB 2; Length 92;

Best Local Similarity 83.3%; Pred. No. 20;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCVOQPG 8

||:||||

Db 82 VCIQPG 87

RESULT 4

US-08-444-792-4
; Sequence 4, Application US/08444792
; Patent No. 5726037
; GENERAL INFORMATION:
; APPLICANT: Bodary, Sarah C.
; APPLICANT: Gorman, John W.
; APPLICANT: McLean, John W.
; APPLICANT: Napier, Mary A.
; TITLE OF INVENTION: METHOD FOR PREPARING WATER SOLUBLE
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,792
; FILING DATE: 19-May-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380227
; FILING DATE: 30-JAN-1995
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/218878
 ; FILING DATE: 28-MAR-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/821337
 ; FILING DATE: 13-JAN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/444490
 ; FILING DATE: 01-DEC-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/290224
 ; FILING DATE: 22-DEC-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lee, Wendy M.
 ; REGISTRATION NUMBER: 00,000
 ; REFERENCE/DOCKET NUMBER: P0552P1C3D4
 ; TELEPHONE: 415/225-1994
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 718 amino acids
 ; TYPE: Amino Acid
 ; TOPOLOGY: Linear
 ;
 ; US-08-444-792-4

Query Match 73.9%; Score 34; DB 1; Length 718;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCVQPG 8
 II:III
 DB 613 VCIQPG 618

RESULT 5
 US-08-445-042-4
 ; Sequence 4, Application US/08445042
 ; Patent No. 5726290
 ; GENERAL INFORMATION:
 ; APPLICANT: Bodary, Sarah C.
 ; APPLICANT: Gorman, Cornelia M.
 ; APPLICANT: McLean, John W.
 ; APPLICANT: Napier, Mary A.
 ; TITLE OF INVENTION: METHOD FOR PREPARING WATER SOLUBLE
 ; POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 720 kb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WinPatIn (Genentech)
 ;
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/445.042
 ; FILING DATE: 19-May-1995
 ; CLASSIFICATION: 530
 ;
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/380227
 ; FILING DATE: 30-JAN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/218878
 ; FILING DATE: 28-MAR-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/821337
 ; FILING DATE: 13-JAN-1992

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/444490
 ; FILING DATE: 01-DEC-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/290224
 ; FILING DATE: 22-DEC-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lee, Wendy M.
 ; REGISTRATION NUMBER: 00,000
 ; REFERENCE/DOCKET NUMBER: P0552P1C3D1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-1994
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 718 amino acids
 ; TYPE: Amino Acid
 ; TOPOLOGY: Linear
 ;
 ; US-08-445-042-4

Query Match 73.9%; Score 34; DB 1; Length 718;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCVQPG 8
 II:III
 DB 613 VCIQPG 618

RESULT 6
 US-07-728-215-32
 ; Sequence 32, Application US/07728215
 ; Patent No. 5962643
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Dean
 ; APPLICANT: Quaranta, Vito
 ; APPLICANT: Pytela, Robert
 ; TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
 ; THEREOF
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
 ; STREET: 4370 La Jolla Village Drive, Suite 700
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 92122
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ;
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/728.215
 ; FILING DATE: 19910711
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Campbell, Cathryn A.
 ; REGISTRATION NUMBER: 31,815
 ; REFERENCE/DOCKET NUMBER: P31 8717
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 535-9001
 ; TELEFAX: (619) 535-8949
 ; INFORMATION FOR SEQ ID NO: 32:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 788 amino acids
 ; TYPE: AMINO ACID
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ;
 ; US-07-728-215-32

Query Match 73.9%; Score 34; DB 2; Length 788;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCVQPG 8
||:| ||
Db 613 VCIQPG 618

RESULT 7
US-08-851-843A-205
; Sequence 205, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-00293005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 205:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-851-843A-205

Query Match 69.6%; Score 32; DB 3; Length 76;
Best Local Similarity 62.5%; Pred. No. 37;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8
||:| ||
Db 49 GRLCAGPG 56

RESULT 8
US-08-974-549A-324
; Sequence 324, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA: WO PCT/US97/17885
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-00261005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 324:

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-549A-324

Query Match 69.6%; Score 32; DB 4; Length 76;
Best Local Similarity 62.5%; Pred. No. 37;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8
||| ||
Db 49 GRLCAGPG 56

RESULT 9
US-08-652-816A-12
; Sequence 12, Application US/08652816A
; Patent No. 5872215
; GENERAL INFORMATION:
; APPLICANT: Osbourn, JK
; APPLICANT: Allen, DJ
; APPLICANT: McCafferty, JG
; TITLE OF INVENTION: Specific binding members, materials and
; TITLE OF INVENTION: methods.
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,816A
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.4
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.8
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206318.9
; FILING DATE: 24-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206372.6
; FILING DATE: 23-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9525004.9
; FILING DATE: 07-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610824.6
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/02240
; FILING DATE: 02-DEC-1992
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/244,597
; FILING DATE: 01-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: David W. Clough
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28111/33308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300

;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 118 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-652-816A-12

Query Match 69.6%; Score 32; DB 2; Length 118;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVQPG 8
|| ||||
Db 8 GRALVQPG 15

RESULT 10
US-08-445-135-2
; Sequence 2, Application US/08445135
; Patent No. 5658789
; GENERAL INFORMATION:
; APPLICANT: Quaranta, Vito
; APPLICANT: Hormia, Marketta
; TITLE OF INVENTION: Promotion of Epithelial Cell Adhesion
; TITLE OF INVENTION: and Hemidesmosome Assembly by a Laminin-Like Molecule
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,135
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/151,134
; FILING DATE: 12-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: DESMOS.002CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; IMMEDIATE SOURCE:
; CLONE: 150 KD
US-08-445-135-2

Query Match 69.6%; Score 32; DB 1; Length 770;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCVPQPG 8

Db 560 VCLQPG 565
||:||||

RESULT 11
US-08-465-380-48
; Sequence 48, Application US/08465380
; Patent No. 5863894
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,380
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma ceylanicum
US-08-465-380-48

Query Match 67.4%; Score 31; DB 2; Length 89;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVCVQPG 8
|||:|
Db 48 RVCFRPG 54

RESULT 12
US-08-486-397-48
; Sequence 48, Application US/08486397
; Patent No. 5866542
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,

; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 357
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,397
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/269
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma ceylanicum
US-08-486-397-48

Query Match 67.4%; Score 31; DB 2; Length 89;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVCVQPG 8
|||:|
Db 48 RVCFRPG 54

RESULT 13
US-08-486-399-48
; Sequence 48, Application US/08486399
; Patent No. 5866543
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles

```

; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,399
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma ceylanicum
;
US-08-486-399-48

```

```

Query Match 67.4%; Score 31; DB 2; Length 89;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 RVCVQPG 8
DB 48 RVCFRPG 54

```

```

RESULT 14
US-08-461-965-48
; Sequence 48, Application US/08461965
; Patent No. 5872098
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansmans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,965

```

```

; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 210/243
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma ceylanicum
;
US-08-461-965-48

```

```

Query Match 67.4%; Score 31; DB 2; Length 89;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 RVCVQPG 8
DB 48 RVCFRPG 54

```

```

RESULT 15
US-08-634-641-48
; Sequence 48, Application US/08634641
; Patent No. 5955294
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George P. Vlasuk
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Mensens, Joris Hilda Lieveen
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansmans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: Suite 4700
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/634,641
; FILING DATE: April 19, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397

```

; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 219/136
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma ceylanicum
; US-08-634-641-48

Query Match 67.4% Score 31; DB 2; Length 89;
Best Local Similarity 71.4% Pred. No. 64;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 RVCVQPG 8
 | | | : | |
Db 48 RVCFRPG 54

Search completed: June 13, 2001, 14:16:31
Job time: 492 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:14:29 ; Search time 229.28 seconds
(without alignments)
1.995 Million cell updates/sec

Title: PCT-US01-05825A-2

Perfect score: 45

Sequence: 1 GRVCVQDG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0401.*

1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	8	21 Y79106	Peptide antagonist
2	42	93.3	8	21 Y79110	Peptide antagonist
3	42	93.3	399	13 R20006	Zonula occludens t
4	38	84.4	8	21 Y79105	Peptide antagonist
5	38	84.4	8	21 Y79114	Peptide antagonist
6	38	84.4	8	21 Y79118	Peptide antagonist
7	35	77.8	8	21 Y79108	Peptide antagonist
8	35	77.8	8	21 Y79109	Peptide antagonist
9	35	77.8	8	21 Y79122	Peptide antagonist
10	34	75.6	782	21 B23609	Human secreted pro
11	34	75.6	785	21 B23636	Human secreted pro

12	73.3	196	20	Y36831	Amino acid sequenc
13	73.3	527	21	Y44560	Human Rhotekin var
14	73.3	544	21	B26790	Human Ras correlat
15	73.3	551	18	W28855	Mouse Rhotekin, bl
16	73.3	563	21	Y44559	Human Rhotekin pro
17	73.3	742	16	R74094	Human zona pelluci
18	73.3	745	15	R55206	Human zona pelluci
19	73.3	745	20	Y42480	Human zona pelluci
20	73.3	745	20	W81817	Human ZPA protein.
21	73.3	745	21	W82215	Human zona pelluci
22	73.3	745	21	Y52689	Human oocyte zona
23	73.3	745	21	Y52180	Human zona pelluci
24	73.3	745	21	Y52984	Human zona pelluci
25	73.3	961	15	R49039	Prpase PTP35 #2.
26	73.3	979	20	Y06606	Human islet cell a
27	71.1	8	21	Y79112	Peptide antagonist
28	71.1	228	20	Y35642	Amino acid sequenc
29	71.1	431	21	B59009	Breast and ovarian
30	71.1	465	21	B07747	A human cancer-ass
31	71.1	949	20	Y37051	Amino acid sequenc
32	68.9	8	21	Y79113	Peptide antagonist
33	68.9	8	21	Y79117	Peptide antagonist
34	68.9	8	21	Y79126	Peptide antagonist
35	68.9	84	20	Y30424	Mature nematode ex
36	68.9	97	21	B40473	Human OREX ORF237
37	68.9	162	20	Y30437	Mature nematode ex
38	68.9	162	21	B15322	A. caninum nematod
39	68.9	181	17	R91712	AcanAP47. Ancylos
40	68.9	181	20	Y30410	Nematode extracted
41	68.9	217	19	W75215	Human secreted pro
42	68.9	260	20	W88312	Sugar transferase
43	68.9	322	22	B65019	Human secreted pro
44	68.9	327	20	Y05529	Soybean vestitone
45	68.9	327	21	Y97832	Soybean vestitone

ALIGNMENTS

RESULT 1
Y79106
ID Y79106 standard; Peptide; 8 AA.
XX
AC Y79106;
XX
DT 05-JUN-2000 (first entry)
XX
DE Peptide antagonist of zonulin.
XX
KW Zonulin; antagonist; zonula occludens toxin receptor;
KW blood-brain barrier; antiinflammatory; cerebroprotective;
KW neuroprotective; dermatological; antiulcer; antiviral;
KW antibacterial; cytostatic; anti-HIV; vulnerary; antiallergic;
KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
KW gastrointestinal inflammation; therapy.
XX
OS Synthetic.
XX
PN WO200007609-A1.
XX
PD 17-FEB-2000.
XX
PF 28-JUL-1999; 99WO-US16683.
XX
PR 03-AUG-1998; 98US-0127815.
XX
PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX
PI Fasano A;
XX
DR WPI; 2000-205565/18.
XX
PT New peptide antagonist of zonulin useful as antiinflammatory agent for

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 XX shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 PS Claim 1; Page 41; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

XX Sequence 8 AA;

Query Match 100.0%; Score 45; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. NO. 3.2e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVQDG 8
 |||
 Db 1 grvcvqdg 8

RESULT 2

ID Y79110 standard; Peptide; 8 AA.

XX AC Y79110;

XX 05-JUN-2000 (first entry)

XX Peptide antagonist of zonulin.

XX Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antiulcer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnary; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.

XX Synthetic.

OS WO200007609-A1.

XX 17-FEB-2000.

XX 28-JUL-1999; 99WO-US16683.

XX 03-AUG-1998; 98US-0127815.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX

PI Fasano A;
 XX WPI; 2000-205365/18.
 XX
 XX New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 PS Claim 1; Page 42; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

XX Sequence 8 AA;

Query Match 93.3%; Score 42; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. NO. 3.2e+05;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVQDG 8
 |||
 Db 1 grvcvqdg 8

RESULT 3

ID R20006 standard; Protein; 399 AA.

XX AC R20006;

XX 31-MAR-1992 (first entry)

XX Zonula occludens toxin.

XX ZOT; cholera; vaccine; enterotoxin; diarrhoea.

XX Vibrio cholerae.

XX WO9118979-A.

XX 12-DEC-1991.

XX 05-JUN-1991; 91WO-US03812.

XX 05-JUN-1990; 90US-0533315.

XX (UYMA-) UNIV MARYLAND BALTI.

XX

PI Kaper JB, Baudry-Maurelli B, Fasano A;
 XX WPI; 1992-007465/01.
 DR N-PSDB; Q20185.
 XX
 PT New *Vibrio cholerae* strains - comprise restriction endonuclease
 PT fragment encoding toxin, used as vaccines against cholera
 XX
 PS Disclosure; Fig 18; 83pp; English.
 XX
 CC The amino acid sequence is that of the zonula occludens toxin (ZOT).
 CC It may be responsible for diarrhoea in some strains of cholera and
 CC the ZOT gene or fragments of it are deleted from strains of *Vibrio*
 CC *cholerae* (V.c.) to be used as vaccines. These V.c. strains have 100%
 CC efficacy in protecting humans against subsequent infection with a
 CC strain of a similar serotype and avoid undesirable side effects such
 CC as diarrhoea, nausea and cramping. Cultures of these strains may be
 CC used for prodn. of vaccines against cholera.
 XX
 SQ Sequence 399 AA;
 Query Match 93.3%; Score 42; DB 13; Length 399;
 Best Local Similarity 87.5%; Pred. No. 5.6;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GRVCVQDG 8
 Db 291 grlcvcqdg 298
 AC Y79105;
 XX
 XX 05-JUN-2000 (first entry)
 XX Peptide antagonist of zonulin.
 XX
 KW Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antiulcer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnerary; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.
 XX
 OS Synthetic.
 XX
 XX WO2000007609-A1.
 XX 17-FEB-2000.
 XX
 PF 28-JUL-1999; 99WO-US16683.
 XX
 PR 03-AUG-1998; 98US-0127815.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Fasano A;
 XX
 DR WPI; 2000-205565/18.
 XX
 XX New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 PS Claim 1; Page 41; 69pp; English.
 XX
 CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not

CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 XX
 SQ Sequence 8 AA;
 Query Match 84.4%; Score 38; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GRVCVQDG 8
 Db 1 grvcvcpg 8
 AC Y79114;
 XX
 XX 05-JUN-2000 (first entry)
 XX Peptide antagonist of zonulin.
 XX
 KW Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antiulcer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnerary; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.
 XX
 OS Synthetic.
 XX
 XX WO2000007609-A1.
 XX 17-FEB-2000.
 XX
 PF 28-JUL-1999; 99WO-US16683.
 XX
 PR 03-AUG-1998; 98US-0127815.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Fasano A;
 XX
 DR WPI; 2000-205565/18.
 XX
 XX New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX

PS Claim 1; Page 44; 69pp; English.

XX This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

XX Sequence 8 AA;

Query Match 84.4%; Score 38; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVQDG 8

Db 1 grgcvcqdg 8

RESULT 6

ID Y79118 standard; Peptide; 8 AA.

AC Y79118;

DT 05-JUN-2000 (first entry)

XX Peptide antagonist of zonulin.

XX Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antiulcer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnary; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.

XX Synthetic.

XX WO200007609-A1.

XX 17-FEB-2000.

XX 28-JUL-1999; 99WO-US16683.

XX 03-AUG-1998; 98US-0127815.

XX (UTMA-) UNIV MARYLAND BALTIMORE.

XX Fasano A;

XX WPI; 2000-205565/18.

XX

PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis

PS Claim 1; Page 45; 69pp; English.

XX This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

XX Sequence 8 AA;

Query Match 84.4%; Score 38; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVQDG 8

Db 1 grgcvcqdg 8

RESULT 7

ID Y79108 standard; Peptide; 8 AA.

AC Y79108;

DT 05-JUN-2000 (first entry)

XX Peptide antagonist of zonulin.

XX Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antiulcer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnary; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.

XX Synthetic.

XX WO200007609-A1.

XX 17-FEB-2000.

XX 28-JUL-1999; 99WO-US16683.

XX 03-AUG-1998; 98US-0127815.

PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX Fasano A;
 XX WPT: 2000-2055565/18.
 DR
 XX
 XX New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 PS Claim 1: Page 42; 69pp; English.
 XX
 CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 XX
 XX Sequence 8 AA;
 SQ
 Query Match 77.8%; Score 35; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GRVCVQDG 8
 Db ||| ||||
 1 grvlvqdg 8
 RESULT 8
 Y79109
 ID Y79109 standard; Peptide; 8 AA.
 AC
 XX Y79109;
 XX
 XX 05-JUN-2000 (first entry)
 XX Peptide antagonist of zonulin.
 DE
 XX Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antiulcer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnerary; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.
 XX Synthetic.
 OS
 XX WO200007609-A1.
 FN
 XX 17-FEB-2000.
 PD

XX 28-JUL-1999; 99WO-US16683.
 PF XX
 XX 03-AUG-1998; 98US-0127815.
 PR XX
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 PA Fasano A;
 XX WPT: 2000-2055565/18.
 DR
 XX
 XX New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 PS Claim 1: Page 42; 69pp; English.
 XX
 CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 XX
 XX Sequence 8 AA;
 SQ
 Query Match 77.8%; Score 35; DB 21; Length 8;
 Best Local Similarity 75.0%; Pred. No. 3.2e+05;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GRVCVQDG 8
 Db ||| ||||
 1 grlcvqg 8
 RESULT 9
 Y79122
 ID Y79122 standard; Peptide; 8 AA.
 AC
 XX Y79122;
 XX
 XX 05-JUN-2000 (first entry)
 XX Peptide antagonist of zonulin.
 DE
 XX Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antiulcer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnerary; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.
 XX

OS Synthetic.
 PN WO200007609-A1.
 XX
 PD 17-FEB-2000.
 XX
 PF -28-JUL-1999; 99WO-US16683.
 XX
 PR 03-AUG-1998; 98US-0127815.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Fasano A;
 XX
 DR WPI; 2000-205565/18.
 XX
 PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 PS Claim 1; Page 46; 69pp; English.
 XX
 CC This present sequence is that of a peptide antagonist of zonulin
 CC (Z), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer, by
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC colitic disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 XX
 SQ Sequence 8 AA;
 Query Match 77.8%; Score 35; DB 21; Length 8;
 Best Local Similarity 75.0%; Pred. No. 3.2e+05;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GRVCVQDG 8
 Db 1 99icvqdg 8
 RESULT 10
 B23609
 ID B23609 standard; Protein; 782 AA.
 XX
 AC B23609;
 XX
 DT 12-JAN-2001 (first entry)
 XX
 DE Human secreted protein SEQ ID NO: 18.
 XX
 KW Human; secreted protein; cytokine; cell proliferation;
 KW nutritional supplement; immune modulation; autoimmune disorder;
 KW

haematopoiesis regulation; tissue growth; haemostasis; inflammation.
 OS Homo sapiens.
 XX
 PD 10..22
 XX
 PF /label= signal_peptide
 XX
 FT 23..782
 XX
 FT /label= mature_protein
 XX
 PN WO200049134-A1.
 XX
 PD 24-AUG-2000.
 XX
 PF 18-FEB-2000; 200WO-US04340.
 XX
 PR 19-FEB-1999; 99US-0120680.
 PR 23-APR-1999; 99US-0298733.
 PR 17-AUG-1999; 99US-0149639.
 PR 23-SEP-1999; 99US-0155686.
 PR 01-OCT-1999; 99US-0157247.
 PR 29-NOV-1999; 99US-0167822.
 PR 29-NOV-1999; 99US-0167823.
 PR 15-FEB-2000; 2000US-0298733.
 XX
 PA (ALPH-) ALPHAGENE INC.
 XX
 PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
 XX
 DR WPI; 2000-549267/50.
 DR N-PSDB; A93109.
 XX
 PT New secreted proteins and polynucleotides encoding them, which are
 PT derived from Homo sapiens, useful for therapy, diagnosis, and research,
 PT as well as nutritional sources or supplements
 XX
 PS Claim 27; Page 250-253; 309pp; English.
 XX
 CC The present sequence is the sequence of a human secreted protein. Its
 CC cDNA was isolated from an adult brain cDNA library. The proteins
 CC and coding sequences of the invention can be used in the isolation of
 CC similar genes and proteins, in the elucidation of their function in vivo,
 CC and to treat a number of conditions. It is possible that they may have
 CC uses as nutritional supplements, as cytokine or cell proliferation
 CC factors, in immune modulation, where they may be used to treat immune and
 CC autoimmune diseases, as haematopoiesis regulators (treating myeloid or
 CC lymphoid cell deficiencies), in the promotion of tissue growth, they may
 CC have chemokine or chemotactic activity, haemostatic or thrombolytic
 CC activity, or anti-inflammatory activity.
 XX
 SQ Sequence 782 AA;
 Query Match 75.6%; Score 34; DB 21; Length 782;
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GRVCVQD 7
 Db 267 grvcvnd 273
 RESULT 11
 B23636
 ID B23636 standard; Protein; 785 AA.
 XX
 AC B23636;
 XX
 DT 12-JAN-2001 (first entry)
 XX
 DE Human secreted protein SEQ ID NO: 92.
 XX
 KW Human secreted protein; cytokine; cell proliferation;
 KW

KW nutritional supplement; Immune modulation; autoimmune disorder;
 KW haematopoiesis regulation; tissue growth; haemostasis; inflammation.
 OS Homo sapiens.

XX WO200049134-A1.
 XX 24-AUG-2000.
 XX 18-FEB-2000; 2000WO-US04340.
 XX 19-FEB-1999; 99US-0120680.
 XX 23-APR-1999; 99US-0298733.
 XX 17-AUG-1999; 99US-0149639.
 XX 23-SEP-1999; 99US-0155686.
 XX 01-OCT-1999; 99US-0157247.
 XX 29-NOV-1999; 99US-0167822.
 XX 29-NOV-1999; 99US-0167823.
 XX 15-FEB-2000; 2000US-0298733.

XX (ALPH-) ALPHAGENE INC.

XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;

XX WPI; 2000-549267/50.

XX New secreted proteins and polynucleotides encoding them, which are
 PT derived from Homo sapiens, useful for therapy, diagnosis, and research,
 PT as well as nutritional sources or supplements

XX Disclosure: Page 298-300; 309pp; English.

XX The present invention is concerned with a number of secreted proteins and
 CC their coding sequences isolated from various human cDNA libraries. The
 CC proteins and coding sequences can be used in the isolation of similar
 CC genes and proteins, in the elucidation of their function in vivo, and to
 CC treat a number of conditions. It is possible that they may have uses as
 CC nutritional supplements, as cytokine or cell proliferation factors, in
 CC immune modulation, where they may be used to treat immune and autoimmune
 CC diseases, as haematopoiesis regulators (treating myeloid or lymphoid cell
 CC deficiencies), in the promotion of tissue growth, they may have chemokine
 CC or chemotactic activity, haemostatic or thrombolytic activity, or
 CC anti-inflammatory activity. No information about sequences B23632-B23645
 CC is given in the specification.

XX Sequence 785 AA;

Query Match 75.6%; Score 34; DB 21; Length 785;
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GRVCVQD 7

Db 270 grvcvnd 276

RESULT 12

Y36831
 ID Y36831 standard; Protein; 196 AA.

XX Y36831;

XX 07-OCT-1999 (first entry)

DE Amino acid sequence of a Chlamydia trachomatis protein.

XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
 KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
 KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
 KW Bartholinitis; pneumonia; venereal lymphogranulomatosis.

XX Chlamydia trachomatis.

XX WO9928475-A2.
 XX 10-JUN-1999.
 XX 27-NOV-1998; 98WO-IB01939.
 XX 04-NOV-1998; 98US-0107077.
 XX 28-NOV-1997; 97FR-0015041.
 XX 17-DEC-1997; 97FR-0016034.
 XX (GEST) GENSET.
 XX Griffais R;
 XX WPI; 1999-371125/31.
 XX Genome sequence of Chlamydia trachomatis
 XX Disclosure: Page 710-711; 1755pp; English.

XX Y36754-Y37949 are encoded by open reading frames (ORFs) of the genome of
 CC Chlamydia trachomatis (see 201425). The polypeptides can be used as
 CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
 CC can also be used to control growth of the microorganism. Chlamydia
 CC trachomatis is responsible for a large number of diseases, e.g. eye
 CC diseases such as conventional trachoma, nonendemic trachoma, as
 CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
 CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
 CC perihepatitis, Bartholinitis; pneumonia; venereal lymphogranulomatosis;
 CC and venereal lymphogranulomatosis. The polypeptides of the invention
 CC may be of use in treating these diseases.

XX Sequence 196 AA;

Query Match 73.3%; Score 33; DB 20; Length 196;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GRVCVQD 8

Db 150 ggycvqdg 157

RESULT 13

Y44560

ID Y44560 standard; Protein; 527 AA.

XX Y44560;

XX 04-APR-2000 (first entry)

XX Human Rhotekin variant protein.

XX Rhotekin; Rho effector; Chromosome 2; target; GTP-activated Rho protein;
 KW susceptibility locus; Parkinson's disease; cytostatic; cardiac; primer;
 KW vaccine; probe; treat; cancer; cardiovascular disorder; brain disorder;
 KW developmental disorder; cytoskeletal-associated disorder; variant; EST;
 KW expressed sequence tag; signal transduction related disease.

XX Homo sapiens.

OS Synthetic.

XX WO9958667-A1.

XX 18-NOV-1999.

XX 13-MAY-1999; 99WO-GB01525.

XX 14-MAY-1998; 98GB-0010409.

XX (SWIK) SMITHKLINE BEECHAM PLC.

XX Duecker KN, Calmels TPG;
 XX WPI: 2000-126368/11.
 DR N-PSDB; Z29899.
 XX
 XX New polypeptides used to identify agonists, antagonists and inhibitors
 PT for use in therapy
 XX
 XX Claim 12d; Page 25-27; 33pp; English.
 PS
 XX The present amino acid sequence is the human Rho-kinase variant protein,
 CC a putative target for Rho. It is derived from an EST and has close
 CC homology to Rho-kinase polypeptide. It is a member of the Rho effector
 CC protein family and is located on chromosome 2, that has a susceptibility
 CC locus for Parkinson's disease. Rho-kinase protein binds specifically to
 CC GTP-activated Rho proteins and interacts with downstream components of
 CC Rho-dependent signalling pathways. It has cytoskeletal and cardiant
 CC activity. This sequence may be used to identify agonists, antagonists or
 CC inhibitors and detect diseases associated with Rho-kinase. The DNA sequence
 CC is useful as a source of primers and probes, while the proteins may be
 CC used as vaccines. Rho-kinase sequences are used to treat diseases, such as
 CC cancer, cardiovascular diseases, brain disorders, developmental
 CC disorders, cytoskeletal-associated disorders and signal transduction
 CC related diseases.
 XX
 XX Sequence 527 AA;
 SQ

Query Match 73.3%; Score 33; DB 21; Length 527;
 Best Local Similarity 71.4%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVOD 7
 Db 83 grvcisd 89
 ||||: |

RESULT 14
 B26790
 ID B26790 standard; Protein; 544 AA.
 XX
 AC B26790;
 XX
 XX 19-JAN-2001 (first entry)
 DT
 XX Human Ras correlative GTP binding kinase protein sequence.
 DE
 DE Human; Ras-related GTP binding kinase.
 KW
 XX Homo sapiens.
 OS
 XX CN1257924-A.
 PN
 XX 28-JUN-2000.
 PD
 XX 21-DEC-1998; 98CN-0125690.
 PF
 XX 21-DEC-1998; 98CN-0125690.
 PR
 XX (UYFU-) UNIV FUDAN.
 PA
 XX Yu L, Fu Q, Liu Q;
 PI
 XX WPI: 2000-544297/50.
 DR N-PSDB; A99328.
 DR
 XX Human Ras correlative GTP binding kinase -
 PT
 XX Claim 1; Page 13; 16pp; Chinese.
 PS
 XX The present invention discloses a new polynucleotide, the polypeptide
 CC coded by the polynucleotide and the process for preparing the polypeptide

CC using the polynucleotide. The polypeptide is a human Ras-related GTP
 CC binding kinase. The present sequence represents the protein of the
 CC invention.
 XX
 XX Sequence 544 AA;
 SQ

Query Match 73.3%; Score 33; DB 21; Length 544;
 Best Local Similarity 71.4%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVOD 7
 Db 100 grvcisd 106
 ||||: |

RESULT 15
 W28865
 ID W28865 standard; Protein; 551 AA.
 XX
 AC W28865;
 XX
 DT 21-NOV-1997 (first entry)
 DT
 XX Mouse Rho-kinase, binds to Rho protein.
 DE
 XX Tumour inhibition; metastasis; antitumour; Rho binding protein;
 KW GTPase; murine; platelet aggregation; inflammation; gene therapy.
 XX
 XX Mus musculus.
 OS
 XX Key Location/Qualifiers
 FH Binding-site 7.89
 FT /label= Rho_binding_site
 FT
 XX JP09183797-A.
 PN
 XX 15-JUL-1997.
 PD
 XX 28-DEC-1995; 95JP-0354328.
 PF
 XX 28-DEC-1995; 95JP-0354328.
 PR
 XX (KIRI) KIRIN BREWERY KK.
 PA
 XX WPI: 1997-410813/38.
 DR N-PSDB; T86255.
 DR
 XX New protein designated Rho-kinase which binds Rho protein - useful to
 PT inhibit or treat tumour formation and metastasis
 PT
 XX Claim 5; Pages 11-12; 18pp; Japanese.
 PS
 XX This sequence is that of a mouse protein, designated Rho-kinase, which
 CC has active Rho protein-binding activity and inhibits activity of Rho
 CC protein GTPase. Rho-kinase is useful as an inhibitor of tumour formation
 CC and metastasis (e.g. by gene therapy using the cDNA coding sequence).
 CC Rho-kinase is also used in a competitive assay for screening for
 CC inhibitors of the Rho-kinase-Rho binding interaction or it can be used
 CC for the elucidation of the mechanism of platelet aggregation and
 CC inflammation.
 XX
 XX Sequence 551 AA;
 SQ

Query Match 73.3%; Score 33; DB 18; Length 551;
 Best Local Similarity 71.4%; Pred. No. 3e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVOD 7
 Db 106 grvcisd 112
 ||||: |